جامعة المنصورة
كلية التجارة

الإحصاء التطبيقي وبحوث العمليات

د/ عبد اللطيف عبد الفتاح أبو العلا
أستاذ الإحصاء التطبيقي

د / إبراهيم محمد مهدي
أستاذ الإحصاء الإكتواري

د / سلطان محمد عبد الحميد
أستاذ ورئيس قسم الإحصاء والتأمين

د / ياسر محمد العدل
مدرس الإحصاء

٢٠٠٣ - ٢٠٠٤

مكتبة الجلاء الجديدة
المنصورة

بسم الله الرحمن الرحيم

وقل ربِ زدني علما

صدق الله العظيم

الجزء الأول

الإحصاء التطبيقي

أ.د/ عبد اللطيف عبد الفتاح أبو العلا

أ.د/ سلطـان محمـد عبـد الحميـد

مقدمة

تــأتى هــذه الطــبعة فى إطار ما سبق أن اصدر من طبعات فى مجال الإحصاء التطبيقى والاستنتاج الإحصائي. ولكن ما يميز هذه الطبعة عن سابقاتها هو الاتجاه إلى تقديم بعض الأساليب الإحصائية التطبيقية فى مجال التنبؤ واتخاذ القرارات فى ظل عدم اليقين.

وبالــرغم مــن محدوديــة حجــم هذه الطبعة إلا أنه روعى فى اختيار الموضوعات التى تناولتها التدرج والتكامل وإلا يكون الإيجاز فى العرض على حســاب الأساســيات ، من أجل ذلك اختير محتوى هذه الطبعة ليشمل الفصول التالية:

الفصل الأول: تحليل التباين وتصميم التجارب .

الفصل الثانى: التصنيف متعدد الاتجاهات.

الفصل الثالث: تحليل الانحدار البسيط.

الفصل الرابع: تحليل الانحدار المتعدد.

الفصل الخامس: الطرق اللامعلمية.

وكــان الهدف من الفصول الأربعة الأولى هو تحقيق اتساع الرؤية عند اتخاذ القرار سواء على أساس تعدد العينات أو تعدد المتغيرات أو بتوفير ظروف متشــابهة للمتغيرات عند بدء التجربة أو تحريرها من أثر عدم التجانس السابق عــند جمع المعلومات. أما الفصل الخامس والأخير ، فقد قدمنا بعض الأساليب اللامعلمية (أو اللابارامترية) فى اتخاذ القرارات ، وهى مجالات وأساليب حديثة ومتــنوعة ولكــن اختيار بعضها وتقديمه فى هذا المرجع – أن أمكن تجاوزاً أن يطلــق عليه ذلك – جاء لتحقيق التناظر والتقابل مع ما سبق تقديمه من أساليب بارامــترية كــأدوات لاتخــاذ القــرارات أو حيــن تحول طبيعة القياسات دون استخدامها.

وأنـــه وإن كان الهدف أساساً من إعداد هذا الكتاب هو الدارس والباحث فـــى المجـــالات الـــتجارية والاقتصــــادية وهو ما قد تعكسه العديد من الأمثلة والتطبـــيقات الـــتى حفلـــت بهـــا هذه الطبعة ، إلا أن الحاجة إلى هذه الأساليب الإحصـــائية للـــدارس والباحث فى مجالات العلوم التطبيقية والإنسانية كالعلوم الزراعية والتربية والاجتماع بل وبعض الدراسات الطبية والصحية لا تقل الحاجا عنها فى مجالات الدراسات والبحوث الإدارية والاقتصادية فليس أيسر من إحلال متغير محل آخر بما يتفق مع مجالات التطبيق. وهذا ما يميز الإحصاء كأسلوب رقمى للقياس والتحليل والتشخيص عن العلوم الأخرى.

ولعلـــنا نكون قد وفقنا بفضل من الله فى تحقيق بعض ما كنا نهدف إليه من إعداد لهذا المرجع المحدود.

والسلام عليكم ورحمة الله وبركاته،

أ.د/ عبد الطيف عبد الفتاح أبو العلا
أ.د/ سلطـــان محـــمد عبـــد الحـــميـــد

المنصورة فى أكتوبر ٢٠٠٣

الفصل الأول

تحليل التباين وتصميم التجارب

Analysis of Variance & Design of Experiments

محتويات الفصل:

أولاً : تحليل التباين:

(١) اختبارات الفروض بشأن تباين مجتمع ما أو عدة مجتمعات ، توزيع كا٢.

(١ -١) اختبارات الفروض بشأن تباين المجتمع σ^2.

(١-٢) اختبارات الفروض بشأن تباين مجتمعين : $\sigma^2_1 = \sigma^2_2$ ، توزيع ف (F).

(٢) اختبارات الفروض بشأن ($\mu_1 - \mu_2$) أو عدة متوسطات .

(٣) تحليل التباين.

(٤) ملاحظات ختامية.

ثانياً : تصميم التجارب:

(١) تعاريف.

(١ -١) التجربة.

(١ -٢) المعالجات.

(١ -٣) وحدة التجربة.

(١ -٤) خطأ التجربة.

(١ -٥) العشوائية.

ثالثاً : التصميم كامل العشوائية

(١) استخدام عدد متساو من وحدات التجربة فى كل معالجة.

(١ -١) النموذج الرياضي.

(١ -٢) النموذج الحسابي والتحليل.

(٢) التصميم كامل العشوائية : عدد غير متساو من وحدات التجربة لكل معالجة.

(٣) العلاقـــة بيـــن التصـــميم كامل العشوائية حيث (ل = ٢) واختبار الفرض : $\mu_1 = \mu_2$

(٤) المقارنات الفردية.

(٥) تعليق ختامي.

تمارين.

الفصل الأول

تحليل التباين وتصميم التجارب

مقدمة: سوف نبدأ بمعالجة كيفية اختبار تباين مجتمع ما σ^2 وكذلك تباينات عدة مجـتمعات مستقلة $\sigma^2_{ر}$ وتقديم أدوات الاختبار التى تستخدم وهى توزيعات كا^2 وكذلك نسبة التباين ف كتمهيد لموضوع تحليل التباين.

أولاً : تحليل التباين
Analysis of Variance

(١) اختبارات الفروض بشأن تباين مجتمع ما أو عدة مجتمعات مستقلة.

تعرضـــا فـــي مرحلة سابقة إلى اختبارات الفروض الإحصائية المتعلقة بالوســط الحسابي لمجتمع ما μ باستخدام أدوات اختبار إحصائية تربط بين μ وتقديرها $\bar{س}$ من عينة احتمالية أو عشوائية وذلك باستخدام ى (التوزيع المعتاد المعيارى) أو ت (توزيع ت) بشرط توافر شروط معينة . وكذلك لاختبار الفرق بين متوسطين ($\mu_1 - \mu_2$) اعتمادا على تقديراتهما $\bar{س}_1$ ، $\bar{س}_2$ باستخدام اختبار ى للتوزيع المعتاد أو اعتماداً على توزيع ت متى توافرت شروط معينة . كما تعرضنا أيضاً لاختبار نسبة متغير ما في المجتمع ل أو الفــرق بيــن نسبتيــن ل$_1$ ، ل$_2$ باستخدام تقديراتهما من عينات عشوائية من المجتمع أو المجتمعين أي $\hat{ل}_1 - \hat{ل}_2$ وذلك باستخدام أدوات اختبار تعتمد على التوزيع المعتاد المعيارى.

ولكـــنه مـــع اختبار الفروض الإحصائية المتعلقة بالوسط أو الفرق بين الوســطين والنســبة والفرق بين نسبتين فإنه قد يعنينا أيضا اختبار درجة تباين ظاهــرة أو متغير ما σ^2 أو مقارنة تباين متغير ما بتباين آخر . فعادة ما يكون اهــتمام وحدة مراقبة وفحص الإنتاج فى مصنع ما – كمصنع للمياه الغازية أو بعض المواد الغذائية السائلة كالزيوت – فقد لا يقتصر الاهتمام فقط على متوسط

حجم المياه الغازية المعبأة فى الزجاجات ، ولكن تباين حجم السائل المعبأ والذى قـد يخـتلف من زجاجة لأخرى لأسباب قد يمكن ضبطها أو عدم إمكان التحكم فيها . هذا التفاوت أو التباين لا يقل أهمية لأثره على المستهلك إقبالا أو إعراضا وعلى الجهات أو الأجهزة المعنية بالرقابة على الإنتاج .

(١-١) اختبارات الفروض بشأن تباين مجتمع ما σ^2 :

مثال (١-١):

إذا رمزنا إلى تباين توزيع حجم المياه الغازية فى الزجاجات فى مصنع مـا بـ σ^2 وتقديرها من العينة التى تسحب دوريا للفحص بـ ع^2 وبفرض أن خطـة الفحـص الدورى تقضى بأن تكون ع < ٠,٠١ وبفرض أن حجم عينة الفحص ن = ١٠ زجاجات وبفرض أن ع = ٠,٠٤ فهل يدل ذلك على أن نتيجة الفحص تقدم دليلا كافيا على أن ع^2 تختلف معنويا عن σ^2 = ٠,٠١ عند مستوى المعنوية ٥% ؟ وللوصول إلى قرار فى هذا الشأن فيمكن أن نعتمد على توزيع احتمالى هو كا^2 (chi-square χ^2) ويعرف فى حالتنا هذه كالآتى :

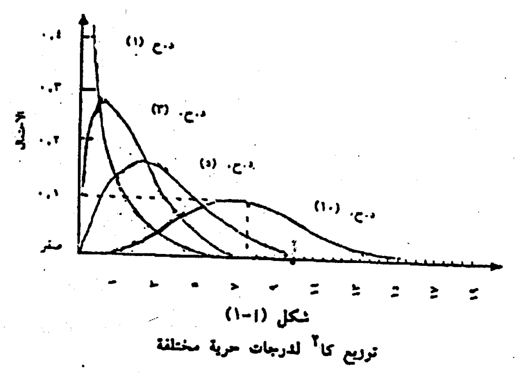
$$\text{كا}^2 = \frac{(\text{ن}-1)\,\text{ع}^2}{\sigma^2} \qquad (1/1)$$

وهو توزيع المعاينة لهذا المتغير ع^2 عند توافر شروط معينة

توزيع كا^2 Chi-square distribution :

وهو توزيع إحتمالى اكتشفه عالمى الإحصاء المعروفين فيشر وبيرسن Sir Ronald Fisher & Karl Pearson وذلـك فى مطلع القرن الماضى . ولهـذا التوزيع مؤشر (بارامتر) واحد هو درجات الحرية (ن-١ فى مثالنا هذا) وبالتالى فهو ليس توزيعا وحيدا كالتوزيع المعتاد المعيارى ، ولكنه مجموعة أو عائلـة من التوزيعات تختلف باختلاف درجات الحرية ، كما أنه توزيع موجب مسـتمر أى يقـع بأكمله على يمين المحور الرأسى ، أى أن جميع قيمة موجبة وهو توزيع موجب الالتواء لدرجات الحرية صغيرة العدد ، ويقترب من التماثل

كلمـا كـبر عـدد تلك الدرجات ، ولهذا التوزيع استخدامات عديدة خاصة فى الاختبارات اللامعلمية كما سنرى فى باب لاحق . والشكل التالى يعرض صورا مختلفة لهذا التوزيع ويقدم جدول رقم (١٠) فى نهاية هذا المرجع الاحتمالات أو المسـاحات تحت منحنى التوزيع لدرجات حرية مختلفة أى أن الذيل الأيمن أو الطرف الأعلى للمنحنى يحدد المساحة ح $(\chi^2 > \chi^2_\alpha) = \alpha$



شكل (١-١)

توزيع كا٢ لدرجات حرية مختلفة

وتطبيقا لذلك على المثال السابق (١-١) حيث ن = ١٠ ، σ^2 = ٠,٠١ ، ع = ٠,٠٤ فإن الاختبار الإحصائى يتم كالآتى :

الفرض العدمى : σ^2 = ٠,٠١ والفرض البديل σ^2 > ٠,٠١

أداة الاختبار الإحصائى كا٢ = $\frac{(ن-1)\ ع^2}{\sigma^2}$

لهـا توزيــع كا٢ بدرجات حرية ن-١ = ٩ ويرفض الفرض العدمى إذا كانت القـيمة المحسوبة لــ كا٢ أكبر من ١٦,٩١٩٠ وهى القيمة الجدولية كا٢ ٩، ٠,٠٥. ولكن :

كا٢٠ = $\frac{9 \times (0.04)^2}{0.01}$ = ١,٤٤

وبالتالى يقبل الفرض العدمى بأن $\sigma^2 < ٠,٠١$ أو $(\sigma < ٠,١)$ باحتمال $\alpha = ٠,٠٥$.

ملحوظة : أجرى هذا الاختبار بفرض أن توزيع حجم المياه الغازية فى الزجاجات فى المصنع (أى المجتمع) يقترب من التوزيع الطبيعى أو المعتاد بغض النظر عن حجم العينة .

(١-٣) اختبار الفروض بشأن تباين مجتمعين : $\sigma_1^2 = \sigma_2^2$

عند مقارنة تباين توزيع قراءات تسجل عن مجتمعين مستقلين ، كنتائج اختبارين للقدرات أو الصلاحية الذى يطبق على عينة من المتقدمين لشغل وظيفة ما أو القراءات التى تسجلها وحدتى قياس (ترمومترين مثلا) فإنه من المهم أن نقارن التباين لأداء المقياسين أو الاختبارين للوصول إلى قرار فى شأن أى من أداة القياس أو الاختبار يمكن استخدامها . وبفرض أن σ_1^2 ، σ_2^2 ترمزان إلى تباينى المجتمعين الأول والثانى ، فإنه لاختبار الفرض بتساوى أو عدم تساوى التباينين أى :

الفرض العدمى : $\frac{\sigma_1^2}{\sigma_2^2} = 1$ أى $\sigma_1^2 = \sigma_2^2$

الفرض البديل : $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$ أى $\sigma_1^2 \neq \sigma_2^2$

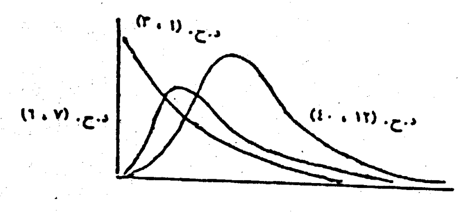
فإن ذلك يتم باستخدام توزيع F-distribution أو توزيع نسبة التباين F Variance ratio حيث :

$$\text{ف} = \frac{\text{ع}_1^2}{\text{ع}_2^2} \qquad (٢/١)$$

توزيع ف Snedecor F-distribution :

وينسب هذا التوزيع الاحتمالى إلى العالم الإحصائى G. W. Snedecor الذى قدمه تكريما للعالم الإحصائى فيشر Fisher الذى عرّف هذه الخاصية باسم "نسبة التباين" وقام بإعداد جداول للعلاقة فى صورة $Z=\log_e \sqrt{F}$ ثم قام Snedecor بتقديم توزيع نسبة التباين فى صورته المجدولة حاليا وهى من الناحية العملية أيسر فى الاستخدام خاصة فى تحليل التباين (جداول من ٤ إلى ٧ فى نهاية هذا المرجع).

وهذا التوزيع ليس توزيعا وحيدا ولكنه توزيع نسبة متغيرين عشوائيين لكل توزيع كا^2 بدرجة حرية $ن_1 - 1$ للبسط ، درجة حرية $ن_2 - 1$ للمتغير فى المقام وبالتالى فهو توزيع يعتمد على معلمتين هما درجات حرية البسط والمقام وبالتالى يختلف باختلاف قيمة هاتين المعلمتين . والشكل التالى يعرض صورا لهذا التوزيع لبعض درجات الحرية .



شكل (١-٢)

توزيع ف لدرجات حرية مختلفة

وتتركز خصائص هذا التوزيع فى الآتى :

١- أنه توزيع موجب .

٢- أنه توزيع موجب الالتواء .

٣- أن منحـــنى الـــتوزيع يقـــترب من المحور الأفقى بازدياد عدد درجات الحرية دون أن يمس المحور .

ويســـتخدم هـــذا التوزيع ، أى توزيع ف لنسبة التباين لاختبار تساوى تبايـــنى مجتمعين مستقلين ، بشرط أن يكون توزيع الظاهرة لكل من المجتمعين توزيـــع معـــتاد وأن تكون العينتين المسحوبتين من المجتمعين عينات عشوائية مســـتقلة كل منهما عن الأخرى . كما يستخدم أيضا لاختبار ومقارنة متوسطات عـــدة مجـــتمعات وهو الأسلوب المعروف باسم "تحليل التباين" (ANOVA) Analysis Of Variance والـــذى سنعرض له فى مرحلة لاحقة متى توافرت شـــروط معيـــنة فـــى توزيعات المجتمعات المقارنة وأهمها أن تكون توزيعات المجـــتمعات المقارنة توزيعات معتادة وأن تكون وحدات القياس من فئة القياس بفـــترة (interval scale) علـــى الأقل وإذا لم يتحقق هذين الشرطين أو ثارت الشـــكوك حول تحققها فتستخدم أحد الأساليب اللامعلمية بدلا من توزيع ف وهو ما سوف نعرض له مستقبلا .

هذا ونذكر أيضا أنه عند استخدام اختبار ت للفرق بين متوسطين $\mu_{١} - \mu_{٢}$ باستخدام أداة الاختبار :

$$ت = \frac{(\bar{س}_{١} - \bar{س}_{٢}) - (\mu_{١} - \mu_{٢})}{ع_{ت}\sqrt{\frac{١}{ن_{١}} + \frac{١}{ن_{٢}}}} \qquad (٣/١)$$

حيث $ع_{ت}$: التباين التجميعي محسوبا من العينتين ، فإنه نفترض تساوى تباين توزيع المجتمعين ، وبالتالي فإنه يتعين استخدام اختبار ف لاختبار تساوى $\sigma_{١}^{٢}$ ، $\sigma_{٢}^{٢}$.

مثال (١-٣):

بفرض أن شركة لسيارات الليموزين لنقل المسافرين من ميدان التحرير إلـى مطـار القاهرة ويمكن أن تستخدم سياراتها أحد مسارين – أى طريقين – ولمـا كـان مـن المهم أن تتحقق الشركة من اتساق وتوافق الوقت المنفق في الوصـول إلـى المطار باستخدام أي من المسارين فقد سجلت الشركة البيانات التالية:

الطريق (أو المسار)	متوسط الوقت بالدقيقة	الانحراف المعياري بالدقيقة	عدد السيارات
الطريق أ	٥٦	١٢	٧
الطريق ب	٥٨	٥	٨

اختبر الفرض بأن $\sigma_1^2 = \sigma_2^2$ عند مستوى المعنوية α = ١٠%.

ويتم تحقيق ذلك كالآتي:

الفرض العدمى : $\sigma_1^2 = \sigma_2^2$ أى $\sigma_1^2 / \sigma_2^2 = 1$

والفـرض البديـل $\sigma_1^2 \neq \sigma_2^2$ أى $\sigma_1^2 / \sigma_2^2 \neq 1$ أى أنه اختبار فى اتجاهيـن ، وبالـتالى فإن مستوى المعنوية يصبح α /٢ = ٥% وهى مساحة الطرف الأعلى (أو الأيمن) من التوزيع وهى القيم الواردة بجدول توزيع ف.

أداة الاختبار : $ف = \frac{ع_1^2}{ع_2^2}$

وبرفض الفرض العدمى إذا كانت ف* (أي المحسوبة من البيانات) أكبر من $ف_{٦، ٧، ٠,٠٥}$ = ٣,٨٧

ولكن ف* = $\frac{١٢^2}{٥^2}$ = ٥,٧٦ > ٣,٨٧ ويرفض الفرض العدمى .

ولكنه من المعتاد أن نحسب ف * = $\frac{ع_1^2}{ع_2^2}$ ، حيث $ع_1^2$ هى الأكبر قيمة (أى $ع_1^2 > ع_2^2$) وفى هذه الحالة فإنه إذا كان الفرض العدمى هو :

$\sigma_1^2 \geq \sigma_2^2$ والفرض البديل هو $\sigma_1^2 > \sigma_2^2$ وعند مستوى المعنوية α = ١٠% وهو اختبار فى اتجاه واحد فيرفض الفرض العدمى إذا كانت

$$ف = \frac{ع_1^2}{ع_2^2} = \frac{١٢^2}{٥^2} = ٥,٧٦ > ف_{٦ ، ٧ ، ١٠\%} = ٢,٨٣$$

وبالتالى فإن نتيجة الاختبار تؤيد قبول الفرض البديل أى أن تباين توزيع الوقت المنفق فى قطع المسافة بأى من المسارين يختلف معنويا عن α = ١٠%.

(٣) الاختبارات الإحصائية المتعلقة بالفرق بين متوسطين والفروق بين عدة متوسطات :

درسنا فى مرحلة سابقة الاختبارات الإحصائية للفرق بين متوسطين ($\mu_1 - \mu_2$) وذلك اعتمادا على توزيع العينات للفرق ($\bar{س}_1 - \bar{س}_2$) سواء باستخدام التوزيع المعتاد المعيارى أو توزيع ت كل بفرض تحقق شروط معينة.

ولكن كثيراً ما تدعوا الحاجة إلى اختبار الفروق بين عدة متوسطات ، فقد تقوم مؤسسة إنتاجية ما إلى تقديم أنواع مختلفة من الحوافز النقدية أو العينية ولتعيين أى من تلك الحوافز أكثر أثرا على الإنتاجية ، فتقوم تلك المؤسسة بتطبيق أنواع الحوافز المختلفة على عينات عشوائية مستقلة من العمال للوصول إلى قرار فى هذا الشأن . فقد تكون الحوافز نقدية بنسبة ثابتة من الأجر أو المرتب بالإضافة إلى ذلك نسبة من الأرباح ، وقد تكون فى صورة عينية أو أجازات بأجر ... ألخ . وفى حالات كهذه وبفرض أننا نقارن بين أثر خمسة أنواع من الحوافز فسوف تكون هناك ١٠ اختبارات ($^{٥}ق_{٢}$) يتعين إجراؤها وهى تتمثل فى اختبار الفروق بين كل متوسطين على حدة وهى: ($\mu_1 - \mu_2$) ، ($\mu_1 - \mu_3$) ، ($\mu_1 - \mu_4$) ، ($\mu_1 - \mu_5$) ، .... ($\mu_4 - \mu_5$).

ومـــن الواضح أن إجراء اختبارات كهذه يعاب عليها الصعوبة العملية التى تنشأ عـــن تعدد الاختبارات ، علاوة على أن تعدد الاختبارات يؤدى إلى زيادة حجم المـــنطقة الحـــرجة عـــن المقرر أصلا قبل بدء الاختبارات ، وهو يعنى الحكم برفض بعض الفروض التى كان يجب أن تقبل لو لم تتكرر المقارنات. وقد تبين أنـــه بـــتكرار اختبار ت عند مستوى المعنوية ٥% للفرق بين متوسطين عدة مـــرات وكانـــت هذه الاختبارات مستقلة ، فإن احتمال الحكم بمعنوية أحد هذه الفروق على الأقل يتجاوز الـــ ٥% ليصل إلى ٢٣% إذا أجرى هذا الاختبار ٥ مرات ويزداد عن ذلك كلما ازداد عدد مرات تكرار الاختبار.

ولقـــد دعى ذلك الاحصائى المعروف سير رونالد فيشر (١٩٣٨) Sir Ronald Fisher إلى تقديم أسلوب تحليل التباين (Analysis of Variance) لتحلـــيل الفـــروق بيـــن المتوسطات حين تتعدد المقارنات . وسوف نعرض فى الفقـــرات التالية لمفهوم عام لهذا الأسلوب ، ثم ننتقل بعد ذلك إلى استخدام هذا الأســـلوب فى تحليل نتائج التجارب التى تصمم بطريقة أو بأخرى لاختبار أثر عامل أو عاملين أو ثلاثة عوامل.

(٣) تحليل التباين (Analysis of Variance):

ســـوف نســـتخدم المثال التالى ببيانات فرضية فى تقديم موضوع تحليل التباين .

مثال (١-٣):

بفـــرض أن البـــيانات التالية تمثل عدد الدقائق التى استغرقت فى كتابة صفحة من تقرير ما على ثلاثة أنواع مختلفة من الآلات الكاتبة (أ ، ب ، جـــ) وقد اختيرت عينة عشوائية من ٩ من الناسخين على الآلة الكاتبة ، و روعى فى اختيارهم التقارب والتجانس فى القدرة على الكتابة على الآلة الكاتبة تم توزيعهم عشوائياً وبالتساوى على الآلات الثلاث ، بحيث خصص منهم ٣ للكتابة على كل من الأنواع الثلاثة من الآلات:

نوع الآلة	عدد الدقائق التى استغرقت فى كتابة الصفحة $س_{رر}$			مجــ س ر	$\bar{س}_{ر}$	مجــ $(س_{رر} - \bar{س}_{ر})^{٢}$
أ	٤	٥	٩	١٨	٦	١٤
ب	٦	٥	٧	١٨	٦	٢
جـ	٨	١٦	١٢	٣٦	١٢	٣٢
				٧٢		٤٨

ويمكنـــنا الوصول إلى تقديرين للتباين لكل عينتين ممكنتين وعددها ٣ أزواج ، أى يمكن إجراء ثلاث مقارنات وذلك على النحو التالى:

التقدير الأول : التباين التجميعى Pooled Variance

$$ع^{٢}_{ت} = \frac{١}{ن_{١} + ن_{٢} - ٢} [مجــ (س - \bar{س}_{١})^{٢} + مجــ (س - \bar{س}_{٢})^{٢}] \quad (١/٤)$$

حيث ترمز ١ ، ٢ فى دليل ن ، $\bar{س}$ إلى العينتين الأولى والثانية (مثلا).

التقدير الثانى: التباين الكلى Total Variance

$$ع^{٢}_{ك} = \frac{١}{ن - ١} \quad مجــ_{ر ر} (س_{ر ر} - \bar{س}..)^{٢} \quad (١/٥)$$

ويعرف هذا التقدير باسم التباين الكلى، حيث ن = ن١ + ن٢ = مجموع مفــردات العينتيــن ، $\bar{س}$.. الوســط الحسابى العام محسوبا من جميع مفردات العينتين معا.

وتطبــيقا للتقديريـــن السابقين وبصرف النظر عن المقام بالاقتصار أى علـــى البســط عـــند حساب مجموع المربعات . فسوف نجد الاتى محسوبا من البيانات السابقة .

بسط التباين التجميعي	بسط التباين الكلى
مجـ ر [مجـ رل (س − $\bar{س}$.ل)^{2}] حيث ر = ١ ، ٢، ٣ ل = أ أو ب أو جـ وهو مجموع المربعات داخل العينات	مجـ (س رل − $\bar{\bar{س}}$..)^{2} ، ن = ن١ + ن٢ وهو مجموع المربعات الكلى
الآلتين أ ،ب :	
مجموع المربعات = ١٤ +٢ = ١٦	$\bar{س}$ أ = $\bar{س}$ ب = $\bar{\bar{س}}$ = ٦ مجموع المربعات = ١٦
الآلتين أ ، جـ :	
مجموع المربعات = ١٤+٢٣=٤٦	$\bar{س}$ أ = ٦ ، $\bar{س}$ جـ = ١٢ ، $\bar{\bar{س}}$ =٩ مجموع المربعات = ١٠٠
الآلتين ب ، جـ :	
مجموع المربعات = ٢+٣٢ = ٣٤	$\bar{س}$ب = ٦ ، $\bar{س}$ جـ = ١٢، $\bar{\bar{س}}$ =٩ مجموع المربعات = ٨٨

ويتضح من النتائج السابقة أن :

مجـ رل (س رل − $\bar{\bar{س}}$..)^{2} − مجـ ر [مجـ ر (س ل − $\bar{س}$.ل)^{2}]

= مجموع المـربعـات الكلـى − مجموع المربعات داخل العينات

بالنسبـة للآلتين أ ، ب = ١٦-١٦ = صفر وكـانت $\bar{س}$ أ = $\bar{س}$ ب

وبالنسـبة للآلتين أ ، جـ = ١٠٠ − ٤٦ = ٥٤ ، $\bar{س}$ أ ≠ $\bar{س}$ جـ

وكذلك في حالة الآلتين ب ، جـ = ٨٨ − ٣٤ = ٥٤ ، $\bar{س}$ ب ≠ $\bar{س}$ جـ

وأخيرا فى حالة الآلات الثلاث أ ، ب ، جـ ، حيث $\bar{\bar{س}}$ = ٨ فإن :

مجـ رل (س رل − $\bar{\bar{س}}$ ..)^{2} = ١٢٠ − ٤٨ = ٧٢ ، $\bar{س}$أ = $\bar{س}$ب ≠ $\bar{س}$جـ

ومما سبق يتبين لنا أنه عندما كان $\bar{س}_م = \bar{س}_ن$ ، م ≠ ن ، فإن :

مجموع المربعات الكلى = مجموع المربعات داخل العينات

فى حين أنه عندما كان $\bar{س}_م \neq \bar{س}_ن$ ، م ≠ ن فإن :

مجموع المربعات الكلى − مجموع المربعات داخل العينات > صفر

= ٥٤ في حالة العينتين لكل من (أ ، جـ) ، (ب ، جـ)

= ٧٢ في حالة العينات الثلاث (أ ، ب ، جـ)

والطـرف الأيسر يشير إلى مجموع المربعات الذي ينشأ بسبب اختلاف المتوسـطات ، أي اختلاف المتوسطات عن المتوسط العام ، ويعرف عادة باسم **مجمـوع المـربعات بين العينات** ، وهذه هي الفكرة من تحليل التباين. وبالتالي فإنه يمكن تعريف أسلوب تحليل التباين بأنه " أسلوب إحصائي يعتمد في اختبار أثر أنواع مختلفة من عامل (أو أكثر) على اختلاف المتوسطات التي تعبر عن اختلاف الأداء باختلاف نوع أو مستوى العامل المستخدم وذلك بتجزئة مجموع المربعات الكلى إلى مركبات يعزى إحداها إلى الاختلاف بين المتوسط محسوبا لكل نوع عن المتوسط العام والمركبة الأخرى إلى الاختلاف داخل كل عينة أي اختلاف مفردات كل عينة عن متوسطها ويعرف باسم البواقي (Residual) أي أنه في حالة التحليل لعامل واحد سنصل إلى النتيجة التالية:

مجمـوع المربعات الكلى = مجموعات المربعات بين العينات (الأنواع) + مجموع المربعات داخل العينات (البواقي).

وبالتالـي فأنـه يمكن تلخيص وتحليل النتائج الكلية للمثال السابق باستخدام أسلوب تحليل التباين على النحو التالي:

(١) مجموع المربعات الكلى الذي يعبر عن التباين الكلى أي أنه يعزى إلى اخـتلاف الوقـت المنفق في كتابة صفحة التقرير بكل من كاتبي الآلة الكاتـبة الـ (٩) أي $س_{رل}$ حيث يرمز ل = أ أو ب أو جـ إلى نوع

الآلة الكاتبة ، ر : ١ ، ٢ ، ٣ إلى الناسخين الثلاث بكل آله كاتبه عن المتوسط العام لوقت الكتابة أيَ $\overline{س}$.. ، أى أن :

مجـــ رل (س رل − $\overline{س}$..)^{٢} = ١٢٠ ويعرف باسم مجموع المربعات الكلى أو م.م.ك.

(٢) مجمـــوع المـــربعات الـــذي يعزى إلى اختلاف متوسط الوقـــت الذي اســـتغرقه كـــــل ناســـخ باستخدام الآلة التي خصصت له من الآلات الثـــلاث أي $\overline{س}$.ر عن المتوسط العام $\overline{س}$.. ويســـــاوى

= ر مجـــ ($\overline{س}$.ر− $\overline{س}$..)^{٢} = ٣(٦−٨)^{٢}+٣(٦−٨)^{٢}+٣(١٢−٨)^{٢}= ٧٢

ويعرف هذا المجموع باسم مجموع المربعات بين أنواع الآلات (م.م.ب)

(٣) مجمـــوع المـــربعات الذى يعزى إلى اختلاف الوقت الذى استنفذه كل ناسخ من الناسخين الثلاث على كل آلة من الأنواع الثلاث عن متوسط الوقت المستنفذ فى الكتابة على الآلة الواحدة من كل نوع ، أى :

مجــــ رل (س رل − $\overline{س}$.ر)^{٢} = ١٤+٢+٣٢ = ٤٨ وهو يساوى الفرق بين م.م.ك ، م.م.ب أى يساوى ١٢٠ − ٧٢ = ٤٨ ولذلك فهو يعرف باسم مجموع المـــربعات المتبقى أو البواقى (Residual) ويرمز له بالرمز (م.م. خ) ويشير إلى التباين العشوائى بين الناسخين على كل آلة أو التباين داخل العينات الثلاث (Within). أى أنه يمكن صياغة هذه العلاقة كالآتى :

مجموع المركبات الكلى (م.م.ك)= مجموع المربعات بين أنواع الآلات (م.م.ب) + مجموع المركبات داخل العينات أو البواقى (م.م.خ) (٦/١)

أى :

S.S. Total = S.S. Between + S.S. Residual

ورقمياً من المثال : ١٢٠ = ٧٢ + ٤٨

ويمكـــن تصوير هذه النتائج فى جدول يعرف باسم جدول تحليل التباين (ANOVA) وهو على الصورة التالية:

جدول تحليل التباين

المصدر	مجموع المربعات (م.م.)	درجات الحرية (د.ح.)	متوسط م.م. (م.م.م)	نسبة التباين (ف)
بين أنواع الآلات (أو بين العينات)	ر مجـ ($\bar{س}_{.ل}$ − $\bar{\bar{س}}$)^٢	(ل − ١)	م.م. بين ÷ (ل −١)	م.م.م.ب / م.م.م.خ
داخل العينات	مجـ ر مجـ ل (س رل− $\bar{س}_{.ل}$)^٢	ل (ر−١)	م.م. داخل÷ ل (ر−١)	
كـلـى	مجـر مجـ ل (س رل− $\bar{\bar{س}}$)^٢	رل − ١	—	—

وتطبيقا على المثال السابق:

المصدر	مجموع المربعات (م.م.)	درجات الحرية (د.ح.)	متوسط م.م. (م.م.م)	نسبة التباين (ف*)
بين العينات	٧٢	٣−١=٢	٣٦	٣٦÷٨ =
داخل العينات **	٤٨	٣(٣−١)=٦	٨	٤,٥
كلى	١٢٠	٩−١=٨		

(** يمكن أن تحسب القيم المناظرة بالفرق)

وتقـارن قـيمة ف * (أى المحسوبة) بقيمة ف لتوزيع ف عند درجات الحرية ٢ ، ٦ ، وعند مستوى المعنوية α المحدد سلفا ويتخذ على أساس مقارنة ف* بـ ف الجدولية يكون القرار بقبول أو رفض الفرض العدمى .

فمثلاً عند α = ٥% وكان الفرض العدمى هو أن متوسط سرعة الأداء على الآلات الثلاث متساو أي $\mu_١ = \mu_٢ = \mu_٣$ ، مقابل الفرض البديل أن الأداء مخـتلف باستخدام نوعين على الأقل من الآلات الكاتبة المستخدمة وعند القيمة الجدولـية ف $_{٢، ٦، ٠,٠٥}$ = ٥,١٤ فإن ف * = ٤,٥ < ٥,١٤ لذلك يقبل الفرض العدمى بعدم اختلاف سرعة الأداء باستخدام أنواع الآلات الثلاث.

وفى الحقيقة فأن الاختبارات تكون للفروق بين المتوسطات $\mu_{ا} - \mu_{جـ}$ ، $\mu_{ب} - \mu_{جـ}$ حيث لا يوجد مبرر لاختبار الفرق من $\mu_{ا} - \mu_{ب}$ لتساوى المتوسطين $\bar{س}_{ا}$ ، $\bar{س}_{ب}$.

ولاختبار الفرض العدمى بأن $\mu_{ا} = \mu_{جـ}$ أى $\mu_{ا} - \mu_{جـ} =$ صفر ، مقابل الفرض البديل بأن $\mu_{ا} \neq \mu_{جـ}$ وباستخدام اختبار ت وهو:

$$ت = \frac{(\bar{س}_{ا} - \bar{س}_{جـ}) - (\mu_{ا} - \mu_{جـ})}{\sqrt{ع^{2}_{ت} \left(\frac{1}{ن_{ا}} + \frac{1}{ن_{ج}}\right)}}$$

$$ت = \frac{١٢-٦}{\sqrt{٤٦ \left(\frac{٢}{٣}\right)}} = - \frac{٦}{٥,٥٣٨} = - ١,٠٨٣$$

وكذلك لاختبار الفرض العدمى $\mu_{ب} - \mu_{جـ} =$ صفر مقابل $\mu_{ب} - \mu_{جـ} \neq$ صفر

نجد أن : $ت = \frac{٦-٢}{\sqrt{٣٤ \left(\frac{٢}{٣}\right)}} = - \frac{٦}{٤,٦٧٠٩} = - ١,٢٨٥$

وبمقارنة ت فى الحالتين بقيمة ت الجدولية عند $\alpha = ٠,٠٥$ ، ٤ درجة حرية ، نجدها = ± ٢,٧٧٦ وكلاهما أصغر من قيمة ت الجدولية ، لذلك يقبل الفرض العدمى فى الحالتين.

وهذه النتائج تتفق مع ما انتهى إليه اختبار الفرض الإحصائي العدمى بأن $\mu_{ا} = \mu_{ب} = \mu_{جـ}$ والذى استخدم فيه اختبار نسبة التباين ف فى الفقرة السابقة.

(٤) ملاحظات ختامية :

(١) يفترض لصحة استخدام أسلوب تحليل التباين لمقارنة متوسطات عينتين أو أكــثر ، أن العيــنات قد سحبت من مجتمعات متساوية فى الأداء مع توافر شروط ثلاث هى :

(١/١) أن يكون توزيع المتغير موضوع القياس فى المجتمعات التى سحبت منها العينات لها التوزيع المعتاد .

(٢/١) وأن σ^2 واحدة لتلك المجتمعات .

(٣/١) أن تكون العينات العشوائية مستقلة.

(٢) يمكــن أن يحســب مجموع المربعات الكلى وكذلك مجموع المربعات بين أنــواع الآلات وداخل العينات ، بطرق جبرية مبسطة غير ما عرض فى الفقرة السابقة ، وهو ما سوف نتناوله بالتفصيل فى مرحلة لاحقة فى تحليل نتائج التجارب.

(٣) يســتخدم أســلوب تحليل التباين فى تحليل الانحدار البسيط والمتعدد وذلك بتجزئه مجموع المربعات الكلى إلى مركبتين (أو أكثر) لتحديد واختبار أثر العلاقــة البسيطة أو المتعددة للمتغير (أو المتغيرات المستقلة) على المتغير الــتابع أو غــير المستقل ، وهو ما سوف يعرضه هذا المرجع عند دراسة تحلــيل الانحدار . وفى الحقيقة فإن أسلوب تحليل التباين وتحليل الانحدار لــيس إلا تطبــيق لأســلوب المــربعات الصغرى Method of Least Squares.

وســوف نــتعرض فى بقية هذا الفصل إلى مدخل فى تصميم التجارب Design of Experiments وســوف نبدأه بمعالجة تصميم وتحليل نتائج التجربة فى اتجاه واحد ، وهو المعروف باسم التصميم كامل العشوائية ، ثم نــنتقل إلى التصميم فى اتجاهين أو أكثر بقدر ما يسمح به مستوى العرض

لهـذا المـرجع ، ويمكن للقارئ متابعة هذا الموضوع بالرجوع إلى بعض المراجع الأكثر تخصصاً الواردة فى نهاية هذا الكتاب.

(٤) وأيـا كان النموذج الذى يتبع فى تصميم وتحليل نتائج التجربة ، فالهدف لا يخرج عن تحقيق واحد أو أكثر من الأهداف التالية:

(٤/١) اختـبار أثـر الأنواع أو المستويات المختلفة لعامل واحد ومعنوية تأثيرها على وحدات التجربة.

(٤/٢) اختـبار متوسـط الأداء لنوع ما أو لمستوى معين للعامل موضوع الـتجربة والدراسة أو اختبار الفرق بين متوسطى (أو متوسطات) الأداء للأنـواع أو المسـتويات المخـتلفة لذلك العامل وكذا تعيين فترات ثقة لمتوسطات الأداء.

(٤/٣) قـياس وتقدير الكفاءة النسبية لنموذج ما مقارنا بنماذج أخرى يمكن استخدامها.

(٤/٤) تقديـر مركـبات التباين للمراحل المختلفة للتجربة ومكوناتها (بين العوامـل وداخل العينات وداخل العينات الفرعية ... إلخ) وهو ما يمكـن الاسـتفادة منه فى إعادة تخطيط النموذج الذى يحقق كفاءه أعلى أو يخفض من نفقات التجربة إذا ما أعيدت فى المستقبل.

ثانياً : تصميم التجارب
Design of Experiments

وسوف نقدم فيما يلى بعض التعاريف السائدة فى هذا المجال:

تعاريف:

(١-١) التجربة Experiment:

انتهـت الفقـرات السـابقة إلـى أنه حين تتعدد المقارنات بسبب تعدد المتغـيرات موضـوع الدراسة أو عدم استقلال بعضها عن البعض ، فإنه تقوم الحاجة إلى تجربة تلك العوامل أو بعضها لقياس أثرها . وعادة ما تقوم التجربة: وهـى استقصـاء مخطـط بهدف الحصول على معلومات مدققه عن متغير ما يعـرف فـى مجـال تحليل الانحـدار باسـم المتغير التابـع أو المتغير غير المسـتقل (Dependent Variable) كمـا يعرف أيضا باسم متغير الاستجابة (Response Variable) فى مجال تصميم التجارب ، وقياس أثر واحد أو أكثر مـن المتغيرات المستقلة عليه وتعرف باسم العوامل (Factors) أو المعالجات (Treatment) أو القطاعـات (Blocks) أو الـتفاعلات (Interactions) فى مجال تصميم التجارب.

ولقد سادت أعمال الإحصائي الكبير R.A.Fisher (١٩١٨ - ١٩٣٨) مجـال تصـميم الـتجارب حين كان مسئولا عن الإحصاء فى محطة التجارب الزراعـية ، الأمـر الـذى أدى إلى أن يسود اسم " المعالجات " " والقطاعات" "والـتفاعلات" "والقطـع" -وهـى مستمدة من المجال الزراعى - لغة تصميم الـتجارب ثـم امتد الأخذ بهذه المسميات فى كافة مجالات تصميم التجارب فى العلـوم الإنسانية والطبية والصيدلية ... إلخ ، وسوف نورد بعض المصطلحات السـائدة فـى هـذا المجال قبل الدخول فى تناول التصميم والتحليل الإحصائي للتجربة ونتائجها.

(١-٢) المعالجات Treatments :

وتطلــق على مستويات العامل موضوع الدراسة ، وحين يكون العامل وحـــيداً كالأجـــر فى صناعة أو عمل ما بحسب النوع (ذكور وإناث) ، فإن هذا العامل يكون ذا مستويين . وحين تشمل التجربة عاملين أو أكثر فإنه يمكن قياس الأثـــر الأساسى لكل عامل بالإضافة إلى الأثر (أو الآثار) المشتركة لمستويات العوامل والتى تعرف باسم تفاعلات.

(١-٣) وحدة التجربة Experimental Unit :

وهـــى الوحـــدة التى تسجل عليها استجابة العامل أو العوامل موضوع الدراسة فى التجربة ، وسيكون مجال دراستنا للتجارب المخططة " Designed Experiment" أى التى يتم فيها ضبط والتحكم فى مواصفات المعالجات وكذلك طريقة تخصيص أو توزيع وحدات التجربة على المعالجات.

فحيـــن يكـــون موضوع الدراسة هو قياس أثر اختلاف وسائل الإعلان على حجم أو قيمة المبيعات من سلعة ما بين وسائل مرئية (الإعلانات بالصحف والـــتلفزيون مـــثلا) أو توزيـــع عينات مجانية عن المنتج لفترة ما ، أو وسائل مســـموعة كالإذاعـــة ، فـــأن طريقة الإعلان عن المنتج هى المعالجات ووحدة التجربة هى قيمة المتفق على الإعلان فى أى من هذه الوسائل .

(١-٤) خطأ التجربة Experimental Error :

ويشير إلى الاختلاف بين الوحدات التجريبية التى تعامل بمعالجة واحدة بســـبب الاخـــتلاف فى طرق المعالجة أو عدم التجانس بين وحدات التجربة أو القصور فى إجراء التجربة.

التكرار وأهميته Replicates: أى تكـــرار الـــتجربة على أكثر من وحدة تجريبية ، وذلك يمكننا من قياس أو تقدير خطأ التجربة وكذلك الارتفاع بدرجة دقـــة الـــتجربة بتصـــغير قيمة ع٢ وكذلك لإمكان التعميم من نتائج التجربة بعد إجراء الاختبارات الإحصائية المناسبة.

(١-٥) العشوائية Randomization:

وهى ضرورية لتحقيق التجانس بين وحدات التجربة ولتأمين الباحث من الوقوع فى خطأ التحيز.

تحليل التباين ANOVA وهو السبيل إلى تحليل نتائج التجربة واختبار الفروض موضوع البحث ، والتعميم لنتائج التجربة من العينة إلى المجتمع الأكبر والأوسع الذى سحبت منه العينة.

وسوف ننتقل إلى معالجة موضوع تصميم التجارب بتقديم بعض النماذج التى تستخدم فى هذا المجال ، وسوف نقتصر على البعض منها وعلى الأخص:

- التصميم كامل العشوائية Completely Randomized Design
- تصميم القطاعات الكاملة العشوائية Complete Randomized Blocks Design
- المربع اللاتينى .
- التحليل العاملى .

وسوف نقدم فى بقية هذا الفصل التصميم أو النموذج كامل العشوائية ثم نقدم بقية التصميمات فى الفصل الثانى ، ويمكن للقارئ أن يرجع إلى مراجع أخرى أكثر تخصصا ذكر البعض منها فى قائمة المراجع فى نهاية هذا المرجع وذلك للتعرف على التصميمات الأخرى.

ثالثاً : التصميم كامل العشوائية
التصنيف فى اتجاه واحد
Completely Randomized Design
One way classification

(١) استخدام عدد متساو من وحدات التجربة في كل معالجة

يستهدف هذا التصميم دراسة أثر أنواع مختلفة لعامل واحد على وحدات تجـربة عشـوائية ، أي بحيـث توزع مفردات التجربة بطريقة عشوائية على الأنـواع المخـتلفة للعامل موضوع الدراسة ، ولذلك يعرف هذا النموذج باسم **التحلـيل لعامل واحد** (One or Single Factor Analysis) كما يعرف أيضاً باسـم **التصنيف أو التحليل فى اتجاه واحد** (One Way Classification) . فإذا كانت الدراسة تستهدف – مثلا – اختبار أثر أنواع مختلفة من الحوافز على إنتاج عمال وحدة إنتاجية ، وبفرض أن أنواع الحوافز ل = ٣ (نقدية ، عينية ، أخـرى) وكانت وحدات التجربة ١٥ عاملاً ، فإنه يمكن توزيع العمال الـ ١٥ علـى الأنـواع الثلاث من الحوافز بطريقة عشوائية ، وقد يتساوى عدد العمال الذيـن يخصصـون لتجربة كل من الأنواع الثلاث أى يختار ر= ٥ من العمال عشوائيا لتجربة أثر كل نوع من الأنواع الثلاث من الحوافز ، وأن لم يكن ذلك ضـرورياً رغم أن تخصيص عدد متساو لكل نوع سوف يؤدى إلى خفض فى الجهد الحسابى.

(١-١) النموذج الرياضى:

وسـواء تسـاوى عـدد مفردات التجربة لكل نوع أو لم يتساوى ، فإن نموذج هذه التجربة سيكون كالتالى:

$$س_{رل} = \mu + مـ_ر + خ_{رل} \qquad (١/٧)$$

حيث ل = ١ ، ٢ ، ... ترمز إلى الأنواع المختلفة للعامل موضوع الاختبار

ويطلق عليها عادة اسم معالجات (Treatments) .

ر = ١ ، ٢ ، ... ن ر عدد وحدات التجربة فى كل معالجة .

μ = المتوسط العام وهو ثابت مجهول القيمة

مـــ ل = وترمز إلى الاثر المضاف للمعالجة ل.

خ رل = الخطأ العشوائى ، ويشير إلى أثر العوامل العشوائية أو التى لا تخضع للقياس أو تكون تحت السيطرة فى التجربة.

س رل : وتشـــير إلـــى ناتج التجربة فى وحدة التجربة أو المفردة ر من المعالجة ل.

ويفترض فى بناء هذا النموذج الآتى:

١- أن نـــاتج الـــتجربة سرل ينشأ عن آثار مضافة (Addiditive) ، هى أثر المعالجـــة ل على وحدة التجربة ر ممثلا فى المتوسط العام μ مضافا إليه أثر المعالجة ل علاوة على تفاوت أو تباين عشوائى .

٢- مـــ ل تعبر عن الأثر المضاف والذى ينشأ بسبب المعالجة ل .

٣- إن الأخطاء خرل تعبر عن أثر العوامل العشوائية وتشمل فى ذلك العوامل الـــتى لا تخضـــع لســـيطرة من يقوم بالتجربة ، ومنها التباين الداخلى بين مفـــردات الـــتجربة ، كما يشمل أثر العوامل الطارئة التى قد لا تتكرر أو تـــتكرر دون انتظام . وسوف يفترض أن لهذه الأخطاء توزيع معتاد توقعه الصـــفر وتباينه σ^2 وإنها مستقلة عن مـــ ل وهو شرط ضرورى لإجراء الاختبارات والتقدير بفترة ثقة .

٤- إن القـــراءات أو القياسات التى تسجل فى التجربة من النوع الكمى لمتغير مستمر وليست من القياسات التصنيفية أو الترتيبية .

ويتعين لتنفيذ تجربة كهذه أن يكون توزيع مفردات التجربة عشوائيا على أنـــواع العامل أو المعالجات ، وبذلك تتاح فرص متساوية لكل معالجة ، كما أن العشـــوائية توفـــر اســـتقلال توزيع الأخطاء وألا تتعرض للارتباط الداخلى أو

الذاتى. ففى التجارب الزراعية مثلا ، يحدث أن تكون القطع المتجاورة متشابهة فــى المواصفات (الخصوبة أو الملوحة) ويعكس ذلك أثره على أثر المعالجات. كمــا أن السكان فى وحدات سكنية متجاورة يتشابهون إلى حد كبير فى الصفات والخصــائص الاجتماعــية والاقتصــادية ، ولذلــك فإن التخصيص العشوائى للمعالجات على الوحدات المتجاورة والمتباعدة يؤدى إلى تلافى التحيز .

(١-٣) النموذج الحسابى والتحليل :

ويمكن تجزئة مجموع المربعات المعرف بالنموذج (٧/١) كالآتى :

$$\text{مجـ}_{\text{رل}} (\text{س}_{\text{رل}} - \overline{\text{س}}_{..})^{2} = \text{ر}\ \text{مجـ}_{\text{ل}} (\overline{\text{س}}_{.\text{ل}} - \overline{\text{س}}_{..})^{2} + \text{مجـ}_{\text{ر}}\,\text{مجـ}_{\text{ل}} (\text{س}_{\text{رل}} - \overline{\text{س}}_{.\text{ل}})^{2} \qquad (٨/١)$$

حيث $\overline{\text{س}}_{.\text{ل}}$: الوسط الحسابى للمشاهدات فى المعالجة ل ،ل = ١ ، ٢ ، ..

$\overline{\text{س}}_{..}$: الوسط الحسابى العام لجميع المشاهدات لجميع المعالجات

$$= \text{مجـ}_{\text{رل}} \frac{\text{س}_{\text{رل}}}{\text{ن}} ،$$

ن : جملة عدد المفردات فى جميع المعالجات فى التجربة.

$\text{س}_{\text{رل}}$: المشاهدة الرائية فى المعالجة ل

أى أن :

مجمــوع المربعات الكلى = مجموع المربعات بين المعالجات (أى الذى يعــزى إلــى الاختلاف بين المعالجات فى الأثر) + مجموع المربعات داخل المعالجات (أو البواقى أو الخطأ العشوائى) (٩/١)

والعلاقة (٨/١) عادة تكتب على النحو التالى كصيغة حسابية :

$$\left(\text{مجـ}_{\text{رل}}\, \text{س}^{2}_{\text{رل}} - \frac{\text{م}^{2}}{\text{ن}}\right) = \left(\text{مجـ}_{\text{ل}} \frac{(\text{م}_{\text{ل}})^{2}}{\text{ن}_{\text{ل}}} - \frac{\text{م}^{2}}{\text{ن}}\right) + \left(\text{مجـ}_{\text{رل}}\, \text{س}^{2}_{\text{رل}} - \text{مجـ}_{\text{ل}} \frac{(\text{م}_{\text{ل}})^{2}}{\text{ن}_{\text{ل}}}\right) \qquad (١٠/١)$$

حيث م : المجموع الكلى لقيم $س_{رل}$ = مجموع المشاهدات

ن : عدد المفردات الكلى فى التجربة .

$م_ل$: مجموع المشاهدات فى المعالجة ل

$ن_ل$: عدد المفردات أو المشاهدات للمعالجة ل = ر فى حالة تساوى عدد الوحدات التجريبية فى كل معالجة .

وعادة ما يحسب الجانب الأيمن من (١٠/١) ثم الحد الأول من الجانب الأيسر ، أما الحد الأخير فيحسب بالفرق أى أن :

مجموع المربعات الكلى − مجموع المربعات بين المعالجات = مجموع المربعات المتبقى (أو الذى يعزى إلى التباين العشوائى) .

ويـــتم تصـــوير هذه النتائج فى جدول خاص يعرف باسم جدول تحليل التبايـــن (Analysis Of Variance) ويختصر عادة بالرمز ANOVA على النحو التالى:

تحليل التباين

المصدر Source	مجموع المربعات (م.م) Sum of Squares (s.s)	درجات الحرية (د.ح) Degrees of Freedom (D.F.)	متوسط المربعات (م.م.م) Mean Squares (M.S.)	نسبة التباين (ف*) Variance Ratio (V.R.)
بين المعالجات Between Treatments	(١) $مجـ_ل \frac{م_ل^٢}{ن_ل} - \frac{م^٢}{ن}$	(٢) (ل − ١)	م.م.م.مـ = (١) ÷ (٢)	
داخل المعالجات (البواقى أو الخطأ العشوائى) Within Treatments or Residual or Error	(٣) بالفرق	(٤) (ن −ل) أو (بالفرق)	م.م.م.خ = (٣) ÷ (٤)	$\frac{م.م.م.مـ}{م.م.م.خ}$
الكلى Total	$مجـ_{رل}\, س_{رل}^٢ - \frac{م^٢}{ن}$	(ن −١)		

الفرض العدمى : $مـ_١ = مـ_٢ = ... = مـ_ل$

والفرض البديل : $مـ_١ \neq مـ_٢ \neq$ .... لأى معالجتين على الأقل .

ولاختبار الفرض السابق ، فإن هذا الفرض يرفض عند مستوى المعنوية α إذا كانت نسبة التباين من العمود الأخير من جدول تحليل التباين أى ف* =

$$\frac{\text{م.م.م}_{ب}}{\text{م.م.م.خ}} > ف_{(ل-١ ، ن-ل ، \alpha)}$$ ويقبل فرض العدم فيما عدا ذلك .

مثال (١-٤):

استخدم أسلوب تحليل التباين (نموذج التحليل فى اتجاه واحد) فى دراسة اخـتلاف إنتاجـية ثلاثة أنواع من الآلات ، وذلك بتوزيع أفراد عينة عشوائية مكونـة من ١٥ عاملاً عشوائيا على كل من الأنواع الثلاث من الآلات ، حيث تمـثل القـراءات المشاهدات المسجلة ، أى عدد الوحدات المنتجة باستخدام هذه الآلات الثلاث فى نهاية اليوم .

الآلة	عدد الوحدات المنتجة س					م ر	$\bar{س}_{ر.}$
أ	٤٧	٥٣	٤٩	٥٠	٤٦	٢٤٥	٤٩
ب	٥٥	٥٤	٥٨	٦١	٥٢	٢٨٠	٥٦
جـ	٥٤	٥٠	٥١	٥١	٤٩	٢٥٥	٥١

مجـ $_{ر ل}$ س$^{٢}_{رل}$ = ٤٠٧٨٤ ، م = ٧٨٠ ، $\bar{س}_{..}$ = ٥٢

وبتحليل نتائج هذه التجربة باستخدام أسلوب تحليل التباين نجد أن :

مجموع المربعات الكلى = مجـ $_{رل}$ س$^{٢}_{رل}$ − $\frac{م^{٢}}{ن}$

$$= ٤٠٧٨٤ - \frac{٧٨٠^{٢}}{١٥} = ٤٠٧٨٤ - ٤٠٥٦٠$$

= ٢٤٤ = م.م.ك

مجمـوع المربعات بين الآلات (أى الذى يعزى إلى اختلاف أثر أنواع الآلات على الإنتاج) = مجـ $_{ر}$ $\frac{م_{ر}^{٢}}{ر}$ − $\frac{م^{٢}}{ن}$ حيث $ن_{١} = ن_{٢} = ن_{٣} = ر = ٥$

$$= \frac{١}{٥}[٢٤٥^{٢} + ٢٨٠^{٢} + ٢٥٥^{٢}] - ٤٠٥٦٠ = ١٣٠ = \text{م.م.ب}$$

وهو فى نفس الوقت = ر مجـــ ر (سَ.ر − سَ..)٢

= ٥ [(٤٩ − ٥٢)٢ + (٥٦ − ٥٢)٢ + (٥١ −٥٢)٢]

= ٥ (٩ + ١٦ + ١) = ١٣٠

= مجمـــوع المـــربعات للفرق بين متوسط الإنتاج لكل نوع من أنواع الآلات والمتوسط العام .

مجمـــوع المـــربعات داخـــل المعالجات (أى مجموع المربعات للخطأ العشوائى) = م.م.ك − م.م.ب

= ٢٢٤ − ١٣٠ = ٩٤ = مجموع المربعات المتبقى

وهو فى نفس الوقت = مجـــ ر [مجـــ ر (س ر ل − سَ.ر)٢]

= [(٤٧ − ٤٩)٢ + (٥٣ − ٤٩)٢ + (٤٩ − ٤٩)٢ + (٥٠ − ٤٩)٢

+ (٤٦ − ٤٩)٢] + [(٥٥ − ٥٦)٢ + (٥٤ − ٥٦)٢ + (٥٨− ٥٦)٢

+ (٦١−٥٦)٢ + (٥٢−٥٦)٢] + [(٥٤ − ٥١)٢ + (٥٠−٥١)٢

+ (٥١−٥١)٢ + (٥١−٥١)٢ + (٤٩−٥١)٢] = ٩٤

= مجمـــوع المـــربعات للفرق بين المشاهدات فى كل معالجة والوسط الحسابى للمعالجة .

ويكون جدول تحليل التباين كالآتى :

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
بين الآلات	١٣٠	٢	٦٥	
البواقى (أو الخطأ العشوائى)	٩٤	١٢	٧,٨٣	٨,٣٠ −
الكلى	٢٢٤	١٤	−	−

الفرض العدمى = مـــ١ = مـــ٢ = مـــ٣ = صفر

والفرض البديل هو أن متوسط أثر المعالجات يختلف لمعالجتين على الأقل .

ويرفض الفرض العدمي حيث :

نسبة التباين ف* = ٨,٣٠ > ف ٢ ، ١٢ ، ٥% = ٣,٨٩

ملحوظة :

كـــان يمكـــن تخفيض الجهد الحسابى بشكل واضح باستبدال قيم س ر ل بالقـــيم س ر ل – و = ح ر ل حيـــث و : وسط فرض مناسب (٥٠ مثلا فى حالتنا هذه) وذلك تطبيقا للقاعدة المعروفة بأن التباين ع^{٢} لا تختلف قيمته العددية سواء حسبت من القراءات الحقيقية أو من القراءات المختزلة بطرح (أو إضافة) مقدار ثابـــت وهى من أسميناه من قبل سواء بالطريقة المباشرة أو بطريقة الفروق (أو الانحرافات) البسيطة .

(٣) التصميم كامل العشوائية (عدد وحدات التجربة لكل معالجة غير متساو) :

ولا يخـــتلف الأمـــر سواء بالنسبة للنموذج الرياضى أو الجبرى سواء كانـــت ن$_{١}$ = ن$_{٢}$ = ... = ن$_{ر}$ أو أنها غير متساوية والمثال التالى يوضح ذلك .

مثال (١–٥) :

لاختبار الفرق بين أثر كل من خمس برامج تدريبية على الآلة الكاتبة ، فقـــد جربت البرامج الخمس على ٢٦ متدربا ، وبعد انتهاء فترة التدريب سجل الوقـــت الذى استغرق فى كتابة تقرير معين لكل من المتدربين الـــ ٢٦ (الوقت مقربا إلى أقرب دقيقة) وكانت النتائج كالآتى :

	الوقت المستغرق فى كتابة التقرير بالدقيقة (س)	م ر
البرنامج أ	٧٨ ٧٢ ٨١ ٨٥	٣١٦
البرنامج ب	٦٣ ٧٤ ٧٥ ٨١ ٧١ ٧٩	٤٤٣
البرنامج جـ	٧٩ ٩٠ ٨٦ ٨١ ٧٤ ٩٥	٥٠٥
البرنامج د	٨٧ ٩١ ٧٥ ٧٩ ٨١ ٨٣ ٩٠	٥٨٦
البرنامج هـ	٧٦ ٧٧ ٨١	٢٣٤
		م=٢٠٨٤

مجـ ر ل س^{٢} ر ل = ١٦٨٣٦٨ ، م^{٢} ÷ ن = ١٦٧٠٤٠,٦١

وتكون نتائج التحليل كالآتى :

مجموع المربعات الكلى = ١٦٨٣٦٨ – ١٦٧٠٤٠,٦١ = ١٣٢٧,٣٩

مجمـــوع المربعات الذى يعزى إلى اختلاف البرامج (أى مجموع المربعات بين البرامج) = $\frac{(٣١٦)^{٢}}{٤} + \frac{(٤٤٣)^{٢}}{٦} + \frac{(٥٠٥)^{٢}}{٦} + \frac{(٥٨٦)^{٢}}{٧}$

$$+ \frac{(٢٣٤)^{٢}}{٣} - ١٦٧٠٤٠,٦١ = ٤٤٤,٢٩$$

أما جدول تحليل التباين فيكون :

جدول تحليل التباين

المصدر	مجموع المربعات م.م	درجات الحرية د.ح	متوسط المربعات م.م.م	نسبة التباين ف*	ف ٥%
بين البرامج	٤٤٤,٢٩	٤	١١١,٠٧		
البواقى (أو الخطأ العشوائى)	٨٨٣,١٠	٢١	٤٢,٠٥	٢,٦٤	٢,٨٤
الكلى	١٣٢٧,٣٩	٢٥	–		

الفرض العدمى : م$_{١}$ = م$_{٢}$ = م$_{٣}$ = م$_{٤}$ = م$_{٥}$

والفرض البديل هو أن الأثر المتوسط لأى برنامجين على الأقل غير متساو .

وهنا يقبل الفرض العدمى ، حيث $ف^{*} = ٢,٦٤ < ف_{٤ ، ٢١ ، ٥\%} = ٢,٨٤$

ملحوظة :

كان يمكن اختصار العمليات الحسابية بشكل واضح فى هذا المثال أيضا بطـــرح وســـط فرضى (٨٠ مثلا) من جميع القراءات وتحليل النتائج باستخدام القـــراءات المختصـــرة وسوف نحصل على نفس النتائج كما فى جدول تحليل التباين الأخير.

(٣) العلاقة بين التصميم كامل العشوائية ، حيث (ل = ٢) واختبار الفرض $\mu_1 = \mu_2$

إذا كانت التجربة العشوائية عبارة عن تجربة أثر عامل واحد يتكون من عامليـــن (ل=٢) فقـــط ، أى تجـــربة أثـــر معالجتين فقط ، فإن اختبار الفرض العدمى:

$م_{١} = م_{٢}$ مقابل الفرض البديل $م_{١} \neq م_{٢}$ ،يقوم على أساس مقارنة نسبة التبايـــن من جدول تحليل التباين بالقيمة المناظرة من جدول توزيع ف بدرجات حرية ١ ، $ن_{١} + ن_{٢} - ٢ = ن - ٢$ عند مستوى المعنوية α ، ويرفض الفرض العدمى أو يقبل فى ضوء نتيجة المقارنة .

وفـــى الحقـــيقة إن هـــذا الاختـــبار لا يختلف فى النتيجة عن اختبار الفرق بين المتوسطين المعروف باختبار ت حيث :

الفرض العدمى : $\mu_{١} = \mu_{٢}$ أو $\mu_{١} - \mu_{٢} =$ صفر

مقابل الفرض البديل $\mu_{١} \neq \mu_{٢}$

حيث يرفض الفرض العدمى متى كانت :

$$ت^{*} = \frac{(\bar{س}_{١} - \bar{س}_{٢}) - (\mu_{١} - \mu_{٢})}{\sqrt{\frac{مجـ_{ر ل}(س_{ر ل} - \bar{س}_{ر .})^{٢}}{ن_{١} + ن_{٢} - ٢}\left(\frac{١}{ن_{١}} + \frac{١}{ن_{٢}}\right)}}$$

$$= \frac{(\bar{س}_1 - \bar{س}_2) - (\mu_1 - \mu_2)}{\sqrt{ع^2_ت \left(\frac{1}{ن_1} + \frac{1}{ن_2}\right)}} \geq ت_{ن_1+ن_2-2 ، 0.02 /\alpha}$$

$$\leq ت_{ن_1+ن_2-2 ، 1-0.02 /\alpha}$$

حيـث $ع^2_ت$ أى التباين التجميعى هى نفسها متوسط مربعات البواقى (أو الخطأ العشوائى) فى جدول تحليل التباين .

أو باختصار $|ت^*| \geq ت_{ن_1+ن_2-2 ، 1-0.02 /\alpha}$ (١١/١)

مثال (١-٦) :

وتطبيقا لذلك فإن اختبار عدم اختلاف إنتاجية الآلتين (أ ، ب) مثلا من بيانات المثال (١-٤) يمكن أن يتم بأحد أسلوبين :

أسلوب تحليل التباين:

مجموع المربعات الكلى = $٤٧^2 + ٥٣^2 + ٤٩^2 + ... + ٥٢^2 - \frac{(٥٢٥)^2}{١٠}$

= ٢٧٧٦٥ – ٢٧٥٦٢,٥ = ٢٠٢,٥

مجموع المربعات بين نوعى الآلة (أ ، ب) أى الذى يعزى إلى اختلاف النوع أ عن ب = $\frac{1}{٥}$ $(٢٤٥^2 + ٢٨٠^2)$ – ٢٧٥٦٢,٥

= ٢٧٦٨٥ – ٢٧٥٦٢,٥ = ١٢٢,٥

إذن مجموع المربعات المتبقى = ٢٠٢,٥ – ١٢٢,٥ = ٨٠

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين	ف ٠,٠١ ٨،١ %
بين الآلات	١٢٢,٥	١	١٢٢,٥	١٢,٢٥	٥,٣٢
المتبقى	٨٠,٠	٨	١٠		
الكلى	٢٠٢,٥	٩			

الفرض العدمى : مـــ ١ = مـــ ٢ = صفر

ويرفض الفرض العدمى حيث $ف_م$ = ١٢,٢٥ > $ف_{١، ٨، ٠٫٠٥}$ = ٥,٣٢

وبأسلوب اختبار الفرق بين متوسطين :

بفرض استقلال العينتين وأن المتغير له توزيع معتاد وأن σ^2 واحدة فى مجتمع الدراسة فإن :

الفرض العدمى : $\mu_1 - \mu_2$ = صفر

الفرض البديل : $\mu_1 \neq \mu_2$

ووسيلة الاختبار هى :

$$|ت_م| = \frac{|(\bar{س}_1 - \bar{س}_2) - (\mu_1 - \mu_2)|}{\sqrt{ع^2_ت \left(\frac{1}{ن_1} + \frac{1}{ن_2}\right)}}$$

ويرفض الفرض العدمى متى كانت $|ت_م|$ > ٢,٣٠٦ حيث $ت_{٨، ٩٧٫٥\%}$ = ٢,٣٠٦

ومن البيانات نجد أن :

$\bar{س}_1$ = ٤٩ ، $\bar{س}_2$ = ٥٦

$$ع^2_ت = \frac{مجـ_ل (مجـ_ر (س_{رل} - \bar{س}_ل)^2)}{ن_1 + ن_2 - 2} \quad ، ل = ١ ، ٢$$

ر = ١ ، ٢ ، ...... ، ن١ أو ن٢

$$= \frac{٥٠ + ٣٠}{٨} = ١٠$$

= متوسط مربعات البواقى فى جدول تحليل التباين الأخير .

$$\therefore |ت_م| = \frac{(٤٩ - ٥٦ - صفر)}{\sqrt{١٠ \left(\frac{1}{5} + \frac{1}{5}\right)}} = \frac{٧}{٢} = ٣,٥$$

وبالتالى يرفض الفرض العدمى .

وفى الواقع فإن :

$(\text{ت}^{*})^{2} = (٣,٥)^{2} = ١٢,٢٥$ = نسبة التباين فى جدول تحليل التباين ، وبالمثل

$\text{ت}^{2}_{\text{٨ ، ٩٧,٥}} = (٢,٣٠٦)^{2} = ٥,٣١٨ = \text{ف}_{\text{١ ، ٨ ، ٥\%}}$

وهذه النتيجة صحيحة لأى قيمة لـــ ف بدرجة حرية واحـــدة فى البسط ، أى أن $\text{ف}_{\text{١ ، ر ، }\alpha} = \text{ت}^{2}_{\text{ر ، }\alpha/٢}$ ، وهى العلاقة بين قيمة ت وقيمة ف الجدولية .

(٤) المقارنات الفردية :

أمـــا وقـــد رفض الفرض العدمى بأن $\text{مـ}_{١} = \text{مـ}_{٢} = \text{مـ}_{٣}$ =صفر فى مثال (٤/١) فإنه قد يعنينا أن نجيب على التساؤل التالى:

أى المعالجات تختلف فى أثرها معنويا ؟ وللإجابة على ذلك فإن المعادلة (٣/١) تصلح لإنشاء العلاقة التالية :

$$|\bar{\text{س}}_{\text{هـ}} - \bar{\text{س}}_{\text{و}}| \leq \text{ت}_{(\text{ن}_{\text{هـ}} + \text{ن}_{\text{و}} - ٢ \text{ ، } \alpha/٢)} \times \sqrt{\text{ع}^{2}_{\text{ت}} \left(\frac{١}{\text{ن}_{\text{هـ}}} + \frac{١}{\text{ن}_{\text{و}}}\right)} \qquad (١٢/١)$$

حيث هـــ ≠ و ، $\text{ن}_{\text{هـ}} = \text{ن}_{\text{و}}$ أو غير متساويان

وتعـــبر العلاقـــة (١٢/١) عن ما يسمى بالحد الأدنى للفروق المعنوية (Least Significant Difference) أو باختصار (LSD) وهى تعنى أن أى فـــرق بين أى متوسطين يساوى أو يتجاوز فى قيمته العددية الطرف الأيسر من تلك العلاقة ، فإنه يعتبر فرقا معنوياً.

وتطبيقاً لذلك على المعالجات التى اختبرت فى المثال (١-٤) نجد أن:

مثال (١-٧):

$$\text{ت}_{\text{٨ ، ٩٧,٥ \%}} \times \sqrt{\text{ع}^{2}_{\text{ت}} \left(\frac{١}{٥} + \frac{١}{٥}\right)} = ٢,٣٠٦ \sqrt{٧,٨٣ \left(\frac{٢}{٥}\right)} = ٤,٨$$

$$|\bar{\text{س}}_{\text{ا}} - \bar{\text{س}}_{\text{ب}}| = ٧ \text{ ، } |\bar{\text{س}}_{\text{ب}} - \bar{\text{س}}_{\text{جـ}}| = ٥ \text{ ، } |\bar{\text{س}}_{\text{ا}} - \bar{\text{س}}_{\text{جـ}}| = ٢$$

والفرقيـــن الأول والـــثانى معنويين والأخير غير معنوى مما يعنى أن وجود ب فى أى من العلاقات الثلاث أدى إلى ظهور الفرق معنوياً ، ومثل هذه النتيجة تفيد فى اتخاذ القرارات.

وجدير بالذكر أننا ما كنا نستطرد فى التحليل للوصول إلى هذه المرحلة لو أننا قبلنا الفرض العدمى بتساوى أثر المعالجات.

وطريقة الحد الأدنى للفروق المعنوية تبدو وكأنها طريقة سهلة وسريعة للكشـــف عن الفروق المعنوية بين أى متوسطين ، ولكن يعاب عليها أنها تؤدى إلـــى زيـــادة مســـاحة منطقة الرفض (α) المقررة سلفا بازدياد عدد المقارنات وهـــناك طرق أخرى تفضل طريقة الحد الأدنى للفروق المعنوية ويمكن للقارئ أن يستشـــير فى ذلك أحد المراجع المتخصصة منها ما ذكر بقائمة المراجع فى نهاية الكتاب .

(٥) تعليق ختامى

(٥-١) يتميز النموذج كامل العشوائية بسهولة التحليل وبتوفير أكبر عدد ممكن من درجات الحرية للخطأ ، كما أن فقد أى مفردة أو مشاهدة – لسبب لا يتصـــل بالتجربة – لا يشكل صعوبة فى التحليل إذ يمكن إسقاطها من الحساب ولا حاجة إلى تقديرها ولا يؤثر ذلك كثيراً على دقة التقديرات ، وأن كان يعاب عليه عدم إمكان التحكم فى الخطأ العشوائى الذى يشمل جميع أنواع التباين عدا ما يرجع إلى المعالجات وهو ما تحاول معالجته النماذج الأخرى مثل نموذج القطاعات العشوائية.

(٥-٢) هـــذا وإذا لـــم يتحقق شرط من شروط صلاحية استخدام نموذج تحليل التبايـــن فـــى حالة النموذج العشوائى الكامل ، فإنه يمكن استخدام أحد أساليب الإحصاء اللامعلمى كاختبار كروسكال – والس والذى سنتعرض له مستقبلاً.

تمارين

١- يعتقد أحد المحللين الماليين أن السهم من النوع أ يحقق عادة سعراً فى التداول فى سوق الأوراق المالية أعلا منه عن السهم من النوع ب . ومع ذلك فإن السهم من النوع أ يمثل مخاطرة أعلى فى التقلبات اليومية للأسعار عنه بالنسبة للسهم من النوع ب وذلك على أساس التباين فى تقلب الأسعار اليومية . وللتحقق من هذا فقد سجلت البيانات التالية عن عينة عشوائية من النوعين من الأسهم وذلك خلال يوم ما:

	السهم من النوع	
	أ	ب
$ن_١ = ن_٢$	٢٥	٢٥
متوسط التقلب فى السعر $\overline{س}$	٠,٢٥٠	٠,١٢٥
ع	٠,٧٦	٠,٤٦

قارن بين درجة تقلب السعر اليومى لنوعى الأسهم أ ، ب وذلك باختبار الفرض العدمى بتساوى درجة تباين تقلب السعر اليومى عند مستوى المعنوية α =٥%.

٢- تعمل شركتان فى تجميع أجهزة التلفزيون . وقد تبين أنه خلال الأيام العشرة الأخيرة من شهر ديسمبر ٢٠٠٢ فقد أرتجع للشركة الأولى عدد ٩ أجهزة فى المتوسط بانحراف معيارى عدد ٢ جهاز . فى حين أنه بالنسبة للشركة الثانية وخلال نفس الفترة كان متوسط عدد الأجهزة المرتجعة عدد ٨,٥ جهاز بانحراف معيارى قدره ١,٥ جهاز.

هل تشير هذه النتائج إلى أن تباين توزيع المرتجع من الأجهزة من انتاج الشركة الأولى أعلى منه للشركة الثانية عند α = ٥% ؟

٣- أ : سجلت قراءات عشوائية عددها ن = ١٠٠ باستخدام آلة قياس وكانت نتيجة هذه التجربة كالآتى: $\overline{س}$ = ٩,٤ وحدة قياس ، ع٢ = ٤,٨٤

اختــبر الفرض العدمى بأن σ^2 = ١ مقابل الفرض البديل σ^2 > ١ عند مستوى المعنوية α = ٠,٠٥

ب : مــاذا لو كان عدد القراءات ن = ٧ ، ع٢ = ٤,٧٤.

فماذا سيكون القرار فى هذه الحالة ؟

٤- الجدول التالى يعرض بيانات جزئية لتحليل التباين:

المصدر	م.م.	د.ح.	م.م.م.	ف
بين الأنواع أو العينات		٢		
داخل العينات أو البواقى			٢٠	
كلى	٥٠٠	١١		

أ : أكمــل جــدول تحليل التباين . كم عدد الأنواع موضوع المقارنة بهذه التجربة ، وما هو حجم العينة الكلى.

ب: حــدد الفــرض موضوع الاختبار (الفرض العدمى) والفرض البديل ثم اختبر الفرض عند α =٥%.

٥- ســحبت عينتين عشوائيتان مستقلتان كل من مجتمع له التوزيع المعتاد لهما المتوسط والتباين (μ_1 ، μ_2) ، (σ_1^2 ، σ_2^2) على التوالى . وكانت النتائج التى سجلت على العينتين كالآتى :

العينة الأولى	العينة الثانية
ن١ = ٢٠	ن٢ = ١٥
$\overline{س}_1$ = ١٢٣	$\overline{س}_2$ = ١١٦
ع١٢ = ٣١,٣	ع٢٢ = ١٢٠,١

أ: اختـــبر الفرض العدمى بأن $\sigma^2_1 = \sigma^2_2$ مقابل الفرض بأن $\sigma^2_1 \neq \sigma^2_2$
استخدم $\alpha = 0.05$.

ب: هل يمكنك استخدام اختبار ت لاختبار الفرض العدمى بأن $(\mu_1 - \mu_2)$ = صفر مقابل الفرض بأن $(\mu_1 - \mu_2) \neq$ صفر . لماذا ؟

٦- الجـــدول الـــتالى يبين عدد الوحدات المنتجة فى خمسة أيام متتالية بواسطة أفراد عينة عشوائية من العمال عددهم ٢٠ عاملاً .

	الأيام					
	١	٢	٣	٤	٥	
	٣٠	٣٠	٤٠	٤٥	٥٥	
	٣٥	٤٠	٤٥	٤٠	٤٥	
	٢٥	٤٥	٥٥	٣٥	٦٠	
	٤٠	٤٥	٣٥	٤٠	٥٠	
المجموع	١٣٠	١٦٠	١٧٥	١٦٠	٢١٠	٨٣٥

استخدم أسلوب تحليل التباين فى تحليل نتائج هذه التجربة عند α = ٥% .

أوجد الخطأ المعيارى للفرق بين أى متوسطين واستخدم ذلك فى إنشاء فترة ثقة للفرق بين أى متوسطين عند α = ٥% وفى ضوء ذلك ناقش الفروق بين متوسطات الإنتاج فى الأيام الخمسة .

٧- البيانات التالية تبين عدد الشيكات التى صرفت خلال كل يوم عمل فى عينة من أربعة فروع لأحد المصارف :

أ	٢٤٥	٢٦٣	٢٦٦	٢٤٩	٣٣٧	٢٧٧	٢٨٩	٢٤٤	٢٦٥		
ب	٢٣٦	١٩١	٢٠٩	٢٥٢	٢١٢	٢٢٤					
جـ	٢٥٣	١٩٢	١٤١	١٦٠	٢٢٩	٢٢١	١٥٠	٢١٥	٢٣٢	٢٣٤	١٩٣
د	١٧٣	١٦٤	١٨٣	١٣٨	١٤٦	١٢٥	١٧٨	١٩٩	١٧٠	١٧٧	١٧٢
	١٨٨	١٩٨									

حلل النتائج السابقة باتباع أسلوب تحليل التباين عند مستوى المعنوية ٥% .

أوجـد مجمـوع المـربعات للخطأ العشوائى ثم بطريقة الجمع من العينات احسب قيمة معامل الاختلاف النسبى حيث معامل الاختلاف النسبى =

$$\frac{ع}{س} \times ١٠٠$$

٨- فيما يلى عدد الساعات التى استغرقت فى إنتاج أربعة أنواع من سلعة مـا (بعد طرح ١٠٠ مائة ساعة) .

السلعة	١	٢	٣	٤	
	٦٤	٧٨	٧٥	٥٥	
	٧٢	٩١	٩٣	٦٦	
	٦٨	٩٧	٧٨	٤٩	
	٧٧	٨٢	٧١	٦٤	
	٥٦	٨٥	٦٣	٧٠	
	٩٥	٧٧	٧٦	٦٨	
م ر.	٤٣٢	٥١٠	٤٥٦	٣٧٢	م = ١٧٧٠
مجـ $س^{٢}_{ر}$	٣١٩٩٤	٤٣٦٥٢	٣٥١٤٤	٢٣٤٠٢	

حلل هذه البيانات مستخدما أسلوب تحليل التباين عند مستوى معنوية ٥% .

٩- البيانات التالية تبين عدد من تغيب عن العمل دون إذن مسبق خلال نصف سنة فى ثلاثة وحدات إنتاجية :

عدد أيام التغيب دون إذن مسبق	الوحدة الإنتاجية			
	أ	ب	جـ	
٢	٦	١	٣	
٣	٤	٣	٥	
٤	٩	٣	٥	
٥	٨	٦	٨	
٦	١	٦	١٩	
٧		١٤	٢٣	
٨		١١	٢٢	
٩		٤	١٤	
١٠		٦	١٤	
١١		٢	٧	
١٢		٣	٨	
١٣		١	٤	
١٤			١	
ر	٣١	٦٠	١٣٣	٢٢٤
م ر.	١٢٥	٤٤٢	١٠٣٧	١٦٠٤
مجـ س^{2}	٥٦١	٣٦٠٢	٨٩٦١	١٣١٢٤

وإذا علم أن ع$^{2}_{1}$ = ١,٩٠ ، ع$^{2}_{2}$ = ٥,٨٦ ، ع$^{2}_{3}$ = ٦,٦٤ هى قيم التباين لتوزيع التغيب فى الوحدات الإنتاجية الثلاثة . أوجد الخطأ المعيارى للفرق بيـن متوسـط طول مدة التغيب فى الوحدتين أ ، ب باستخدام تقدير التباين التجميعى ع$^{2}_{ن}$ ثم باستخدام أسلوب تحليل التباين .

احسب نفس القيمة وعلل اختلاف القيمتين إن وجد . ثم قدر الفرق بين متوسطى مدة التغيب فى الوحدتين بفترة ثقة ٩٥% بالطريقتين .

١٠- فى دراسة عن درجة اختلاف الأسر ذات الحجم الواحد من سكان المناطق السكنية (حضر – ريفية محضرة – ريفية) فى الإنفاق على المأكل والمشرب خلال فترة الدراسة مخفضا بـ ٢٥٠ج فكانت النتائج كالآتى :

ح = س – ٢٥٠

مناطق حضرية	٨١	١٥١	١٠١	٩٦	١١٤
مناطق ريفية محضرة	١٢	٢٥	٢٢	١٩	
مناطق ريفية	٤١-	١٨	٩-		

هل تؤيد هذه النتائج الاعتقاد بعدم اختلاف الأسر فى إنفاقها على المأكل والمشرب باختلاف محل إقامتها عند مستوى المعنوية ٥% ؟

١١- استخدم أسلوب تحليل التباين فى اختبار أثر مدة الخبرة على إنتاجية العامل يوم / جنيه من البيانات التالية :

عدد سنوات الخبرة	عدد العمال فى كل مجموعة	س : الإنتاجية يوم / جنيه				
٨	٥	٣٣	٣٧	٣٢	٣١	٢٧
١٢	٣	٣٠	٣٥	٢٨		
١٦	٤	٣٣	٢٩	٣٤	٣٦	

استخدم وسطا فرضيا و = ٣٠ فى تبسيط عملياتك الحسابية .

الفصل الثانى

التصنيف متعدد الاتجاهات

محتويات الفصل

الفصل الثانى

التصنيف متعدد الاتجاهات

Multiway Classification

مقدمة:

كـــان المصـــدر الوحيد للتباين فى قيم وحدات التجربة هى المعالجات، ولذلـــك فقد تعرضنا فى الباب السابق إلى تصنيف المعالجات بحسب أنواعها أو مســـتوياتها وبالـــتالى كانـــت الخاصية التى تحكم التصميم تفرض اتجاهاً واحداً لـــتوزيع المعالجات سواء كان هذا الاتجاه أفقياً (أى فى شكل صفوف) أو رأسيا (أى أعمـــدة). ثـــم يـــتم توزيـــع وحدات التجربة عشوائياً على جميع أنواع أو مستويات المعالجات، أى أن أسلوب تجميع وحدات التجربة فى مجموعات – كل تمـــثل نوعـــاً أو مســـتوى للمعالجـــة – هو التجمع أحادى العنصر (Single Grouping).

وننتقل الآن إلى توزيع المعالجات بحسب خاصية أخرى، سنطلق عليها اســـم القطاعـــات (Blocks). ونقـــدم فى هذا الصدد نموذج القطاعات الكاملة العشوائية Randomized Complete Blocks Design. ثم ننتقل بعد ذلك إلى توزيـــع المعالجـــات بحســـب خاصـــيتين تخصص لأحدهما الصفوف والأعمدة للخاصية الثانية وهو التصميم المعروف باسم (Latin Squre) وكلاهما يدخل فى إطار النماذج متعددة التصنيفات أو الاتجاهات (Multiway Classification). ثم نختتم معالجاتنا لهذا الموضوع بتقديم محدود لنموذج التحليل العاملى ، نقتصر فيه على نموذج التحليل العاملى ٢×٢.

أولاً : تصميم القاطاعات الكاملة العشوائية
Randomized Complete Blocks Design

(١) القطاعات الكاملة العشوائية بدون تكرار:

لنفرض أن لدينا (ر ل) وحدة تجريبية قسمت إلى ر قطاع (Block) بكل ل معالجـــة (Treatment) وإن كل معالجة خصصت عشوائيا لكل من الوحدات التجريبيـــية فـــى كـــل قطـــاع ، فإن تجربة كهذه تعرف باسم القطاعات الكاملة العشـــوائية. وفى مثل هذا النموذج يكون كل قطاع متجانساً قدر الإمكان، ولكن تتبايـــن القطاعات فيما بينها. ويكون عدد الوحدات التجريبية التى تخصص لكل معالجة فى كل قطاع متساوية فى العدد (عادة وحدة واحدة فى كل معالجة فى كل قطـــاع) ومـــتى تحقـــق ذلـــك فإن الفروق بين القطاعات تكون متعامدة مع أثر المعالجات. ويتبع هذا الأسلوب عادة لاستبعاد أثر عامل ثان هو القطاعات ، هذا العامل قد يكون له أثره بالإضافة إلى العامل الأول وهو المعالجات.

ففـــى مـــثال (١ -٢) فى الفصل السابق ، يمكن أن تكون الأعمدة تمثل العمـــل فى خمسة أيام متتالية أو فى خمسة وحدات مختلفة ، تعبر عن القطاعات فـــى حيـــن تعـــبر الصفوف عن أنواع الآلات الثلاثة أى أن ناتج التجربة يكون كالآتى:

المعالجات (الآلة)	القطاعات (الوحدات الإنتاجية) أ	ب	جـ	د	هـ	م ل
(١)	٤٧	٥٣	٤٩	٥٠	٤٦	٢٤٥
(٢)	٥٥	٥٤	٥٨	٦١	٥٢	٢٨٠
(٣)	٥٤	٥٠	٥١	٥١	٤٩	٢٥٥
م ر	١٥٦	١٥٧	١٥٨	١٦٢	١٤٧	٧٨٠

وكما يتبين من هذا الجدول ، فإن كل معالجة تظهر فى جميع القطاعات وبنفس العدد كما أن كل قطاع يشتمل على جميع المعالجات بنفس عدد الوحدات التجريبية (وحدة واحدة فى كل قطاع فى كل معالجة فى مثالنا هذا).

(١-١) النموذج الرياضى أو النموذج الجبرى (أو الحسابى):

ويمثل هذا النموذج بالمعادلة التالية:

$$س_{رل} = \mu + م_ل + ق_ر + خ_{رل} \qquad (١/٢)$$

ويفترض فى هذا النموذج تخصيص وحدة تجريبية واحدة عشوائيا لكل قطاع وكل معالجة، أى دون تكرار (Single replication) والفروض الخاصة بالنموذج (١/٢) لا تختلف فى شئ عن فروض النموذج (١/١) إلا فى إضافة الحد $ق_ر$ للدلالة على الآثار المضافة للقطاع ر (ر = ١ ، ٢ ، ... ، ل).

ويتكون مجموع المربعات الكلى من مركبات ثلاث ، المركبة الأولى تعبر عن أثر المعالجات والمركبة الثانية تعبر عن أثر القطاعات أما الثالثة فتشير إلى الخطاء العشوائي (أو البواقي) أى التباين الذى لا يفسر على أساس من المعالجات أو القطاعات ، ولكن يرجع إلى عامل الصدفة أو عوامل لا تتكرر بانتظام ويصعب التحكم فيها. أى أن:

مجموع المربعات الكلى = مجموع المربعات بين المعالجات (أو الذى يعزى إلى اختلاف أثر المعالجات) + مجموع المربعات بين القطاعات (أو الذى يعزى إلى اختلاف أثر القطاعات) + مجموع المربعات المتبقى (ويعزى إلى الخطاء العشوائى). (٢/٢)

أى أن:

$$\sum_{\text{رل}} (\text{س}_{\text{رل}} - \bar{\text{س}}_{..})^{2} = \text{ر} \sum_{\text{ل}} (\bar{\text{س}}_{.\text{ل}} - \bar{\text{س}}_{..})^{2} + \text{ل} \sum_{\text{ر}} (\bar{\text{س}}_{\text{ر}.} - \bar{\text{س}}_{..})^{2} + \sum_{\text{رل}} (\text{س}_{\text{رل}} - \bar{\text{س}}_{.\text{ل}} - \bar{\text{س}}_{\text{ر}.} + \bar{\text{س}}_{..})^{2} \qquad (٣/٢)$$

وعادة ما تستخدم الصيغة التالية لحساب مجموع المربعات:

$$\sum_{\text{رل}} \text{س}^{2}_{\text{رل}} - \frac{(\text{م})^{2}}{\text{ر} \times \text{ل}} = \left(\sum_{\text{ل}} \frac{(\text{م}_{\text{ل}})^{2}}{\text{ر}} - \frac{(\text{م})^{2}}{\text{ر} \times \text{ل}} \right) + \left(\sum_{\text{ر}} \frac{(\text{م}_{\text{ر}})^{2}}{\text{ل}} - \frac{(\text{م})^{2}}{\text{ر} \times \text{ل}} \right) + \text{البواقى} \qquad (٤/٢)$$

حيث م : مجموع قيم الوحدات التجريبية = $\sum_{\text{رل}} \text{س}_{\text{رل}}$

$\bar{\text{س}}_{..}$ المتوسط العام = $\frac{1}{\text{ر} \times \text{ل}} \sum_{\text{رل}} \text{س}_{\text{رل}} = \frac{\text{م}}{\text{ر} \times \text{ل}}$

$\text{م}_{\text{ل}}$: مجمـــــوع قيم الوحدات التجريبية للمعالجة ل.

$\bar{\text{س}}_{.\text{ل}}$: متوســـط قيــم الوحدات التجريبية للمعالجة ل.

$\text{م}_{\text{ر}}$: مجمـــــوع قيم الوحدات التجريبية للقـــطاع ر.

$\bar{\text{س}}_{\text{ر}.}$: متوســـــط قيم الوحدات التجريبية للقـــطاع ر.

ويستكمل تحليل النتائج باستخدام جدول تحليل التباين كالآتى:

المصدر	مجموع المربعات (م.م)	درجات الحرية (د.ح)	متوسط المربعات م . م . م	نسبة التباين ف *
بين المعالجات	$\frac{\text{١}}{\text{ر}} \sum_{\text{ل}} (\text{م}_{\text{ل}})^{\text{٢}} - \frac{(\text{م})^{\text{٢}}}{\text{ر} \times \text{ل}}$	ل - ١	$\frac{\text{م.م.مـ}}{\text{ل} - \text{١}}$ = (١)	(١) ÷ (٣)
بين القطاعات	$\frac{\text{١}}{\text{ل}} \sum_{\text{ر}} (\text{م}_{\text{ر}})^{\text{٢}} - \frac{(\text{م})^{\text{٢}}}{\text{ر} \times \text{ل}}$	ر - ١	$\frac{\text{م.م.ق}}{\text{ر} - \text{١}}$ = (٢)	(٢) ÷ (٣)
البواقى	بالفــرق	(ر-١) (ل-١) (أو بالفرق)	$\frac{\text{م.م.خ}}{(\text{ر}-\text{١})(\text{ل}-\text{١})}$ = (٣)	–
الكـــلى	$\sum_{\text{ر ل}} \text{س}^{\text{٢}}_{\text{ر ل}} - \frac{(\text{م})^{\text{٢}}}{\text{ز} \times \text{ل}}$	ر ل - ١	–	–

أولاً : الفرض العدمى $\bar{\text{م}}_{\text{١}} = \bar{\text{م}}_{\text{٢}} = \ldots\ldots = \bar{\text{م}}_{\text{ل}} =$ صفر

والفرض البديل هو اختلاف الأثر المتوسط لأى معالجتين على الأقل ، ويرفض الفرض العدمى إذا كانت:

$$\text{ف}_{\text{م}} = \frac{\text{م.م.مـ}}{\text{ل}-\text{١}} \div \frac{\text{م.م.خ}}{(\text{ر}-\text{١})(\text{ل}-\text{١})} \geq \text{ف}_{(\text{ل}-\text{١})،(\text{ر}-\text{١})(\text{ل}-\text{١})،\alpha}$$

ويقبل الفرض العدمى فى غير ذلك من الأحوال.

ثانياً: الفرض العدمى $ق_1 = ق_2 = ... ق_ر$ = صفر

والفــرض الــبديل هو اختلاف الأثر المتوسط لأى قطاعين على الأقل ويرفض الفرض العدمى إذا كانت:

$$ف_م = \frac{م.م.ق}{ر-1} \div \frac{م.م.خ}{(ر-1)\,(ل-1)} \geq ف_{(ر-1)\,،\,(ر-1)(ل-1)\,،\,\alpha}$$

ويقبل الفرض العدمى فى غير ذلك من الأحوال.

حيث ترمز م.م. مـــ إلى مجموع المربعات للمعالجات، م.م.ق إلى مجموع المربعات للقطاعات أما م.م.خ فتشير إلى مجموع المربعات للخطأ العشوائى أو البواقى .

وسوف نستخدم بيانات الجدول السابق فى عرض طريقة تحليل النتائج لتجربة قطاعات عشوائية كاملة بدون تكرار.

مثال (٢-١)

	الوحدات الإنتاجية						
الآلات	أ	ب	جـ	د	هـ	م ل	$\bar{س}_{.ل}$
(١)	٤٧	٥٣	٤٩	٥٠	٤٦	٢٤٥	٤٩
(٢)	٥٥	٥٤	٥٨	٦١	٥٢	٢٨٠	٥٦
(٣)	٥٤	٥٠	٥١	٥١	٤٩	٢٥٥	٥١
$م_ر$	١٥٦	١٥٧	١٥٨	١٦٢	١٤٧	٧٨٠	٥٢
$\bar{س}_{ر.}$	٥٢	٥٢,٣٣	٥٢,٦٧	٥٤	٤٩	م	$\bar{س}_{..}$

وبفــرض أن بيانات هذا الجدول تمثل إنتاج أفراد عينة عشوائية مكونة من ١٥ عاملا ، وزعوا عشوائيا على وحدات إنتاجية عددها خمسة ، وبكل ثلاثة أنواع مختلفة من الآلات فإن تحليل التباين يظهر التالى:

مجموع المربعات الكلى (م.م.ك) $= \text{مجـ}_{\text{رل}}\ \text{س}^{2}_{\text{رل}} - \dfrac{(\text{م})^{2}}{\text{ر} \times \text{ل}}$

$$= ٤٧^{2} + ٥٣^{2} + \ldots + ٤٩^{2} - \frac{٧٨٠^{2}}{١٥}$$

$$= ٤٠٧٨٤ - ٤٠٥٦٠ = ٢٢٤$$

مجموع المربعات بين الآلات (الذى يعزى إلى اختلاف الآلات)

$$= \text{ر}\ \text{مجـ}_{\text{ل}}\ (\bar{\text{س}}_{\text{ل.}} - \bar{\text{س}}_{..})^{2}$$

$$= ٥\left[(٤٩ - ٥٢)^{2} + (٥٦ - ٥٢)^{2} + (٥١ - ٥٢)^{2}\right]$$

$$= ٥ \times ٢٦ = ١٣٠$$

وهى أيضاً $= \dfrac{١}{\text{ر}}\left[\text{مجـ}_{\text{ل}}\ (\text{م}_{\text{ل}})^{2}\right] - \dfrac{(\text{م})^{2}}{\text{ر} \times \text{ل}}$

$$= \frac{١}{٥}\left[(٢٤٥^{2} + ٢٨٠^{2} + ٢٥٥^{2})\right] - \frac{٧٨٠^{2}}{١٥}$$

$$= \frac{٦٠٠٢٥ + ٧٨٤٠٠ + ٦٥٠٢٥}{٥} - ٤٠٥٦٠$$

$$= ٤٠٦٩٠ - ٤٠٥٦٠ = ١٣٠$$

مجمــوع المــربعات بين الوحدات الإنتاجية (الذى يعزى إلى اختلاف الوحدات الإنتاجية) $= \text{ل}\ \text{مجـ}_{\text{ر}}\ (\bar{\text{س}}_{\text{ر.}} - \bar{\text{س}}_{..})^{2}$

$$= ٣ \times \left[(٥٢ - ٥٢)^{2} + (٥٢,٣٣ - ٥٢)^{2} + (٥٢,٦٧ - ٥٢)^{2} + (٥٤ - ٥٢)^{2} + (٤٩ - ٥٢)^{2}\right] = ٣ \times ١٣,٥٥٧٨$$

$$= ٤٠,٦٧٣٤$$

وهى أيضاً $= \frac{\text{١}}{\text{ل}} \left[\sum_{\text{ر}} (\text{م}_{\text{ر}})^{\text{٢}} \right] - \frac{(\text{م})^{\text{٢}}}{\text{ر} \times \text{ل}}$

$$= \frac{\text{١}}{\text{٣}} \left[\left(\text{١٥٦}^{\text{٢}} + \text{١٥٧}^{\text{٢}} + \text{١٥٨}^{\text{٢}} + \text{١٦٢}^{\text{٢}} + \text{١٤٧}^{\text{٢}} \right) \right] - \frac{\text{٧٨٠}^{\text{٢}}}{\text{١٥}}$$

$$= \frac{\text{١}}{\text{٣}} (\text{٢٤٣٣٦} + \text{٢٤٦٤٩} + \text{٢٤٩٦٤} + \text{٢٦٢٤٤} + \text{٢١٦٠٩})$$

$$- \text{٤٠٥٦٠} = \text{٤٠,٦٧٣٤}$$

ويكون مجموع المربعات المتبقى ويعزى إلى الخطأ العشوائى هو:

= م.م.ك. – م.م. بين الآلات – م.م. بين الوحدات الإنتاجية.

أما جدول تحليل التباين فيكون كالآتى:

المصـــــــــدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
بـــــــين الآلات	١٣٠,٠٠	٢	٦٥	٩,٧٥
بين الوحدات الإنتاجية	٤٠,٦٧	٤	١٠,١٧	١,٥٢
البواقـــــــــى	٥٣,٣٣	٨	٦,٦٧	
الكلـــــــــى	٢٢٤,٠٠	١٤	—	—

الفــرض العدمى : $\bar{\text{م}}_{\text{١}} = \bar{\text{م}}_{\text{٢}} = \bar{\text{م}}_{\text{٣}} =$ صفر

والفرض البديل: $\bar{\text{م}}_{\text{١}} \neq \bar{\text{م}}_{\text{٢}} \neq \bar{\text{م}}_{\text{٣}}$ لأى معالجين على الأقل.

ويــرفض الفرض عند α = ٥ % حيث $\text{ف}_{\text{٢، ٨، ٥\%}}$ = ٤,٤٦ أى أنه لاختلاف أنواع الآلات أثره انمعنوى على وحدات التجربة عند مستوى انمعنوية ٥ % .

أما الفرض العدمى : $\text{ق}_{١} = \text{ق}_{٢} = \text{ق}_{٣} = \text{ق}_{٤} = \text{ق}_{٥} = \text{صفر}$

والفرض البديل: $\text{ق}_{\text{هـ}} \neq \text{ق}_{\text{و}}$ ، هـ ≠ و لأى قطاعين على الأقــل .

وهــنا يقبل الفرض البديل حيث $\text{ف}_{\text{م}} = ١,٥٢ < ٣,٨٤ = \text{ف}_{٤ ، ٨ ، ٥\%}$ ، أى أن اختلاف الوحدات الإنتاجية ليس لها أثر على الوحدات التجريبية.

وفى حالة كهذه ، يمكن إعادة إدماج مجموع المربعات ودرجات الحرية للوحــدات الإنتاجية (طالما أنه تبين عدم معنوية أثرها) إلى البواقى، وبذلك نعود إلى التحليل فى اتجاه واحد ، حيث تتضح عدم جدوى اختلاف الوحدات الإنتاجية فى تفسير جزء من اختلاف الوحدات التجريبية فى القيمة.

(١ -٣) طبيعة حد البواقى:

ســبق أن عرّف مجموع المربعات للخطأ العشوائى أو مجموع المربعات للبواقى بأنه:

$$= \sum_{\text{رل}} (\text{س}_{\text{رل}} - \bar{\text{س}}_{\text{ل.}} - \bar{\text{س}}_{\text{.ر}} + \bar{\text{س}}_{..})^{٢} \qquad (٢/٥)$$

حيث $\bar{\text{س}}_{\text{ل.}}$ ، $\bar{\text{س}}_{\text{.ر}}$ ، $\bar{\text{س}}_{..}$ هى متوسطات المعالجة ل والقطاع ر والمتوسط العــام على التوالى ، ومجموع المربعات طبقاً لهذا التعريف الذى يقدر باستخدام طـريقة المــربعات الصغرى ، يؤدى إلى الحصول على أقل مجموع لمربعات الخطأ الذى يحقق أحسن تقدير لـــ $\text{مـ}_{\text{ل}}$ ، $\text{ق}_{\text{ر}}$ ، μ وهى معالم النموذج:

$$\text{س}_{\text{رل}} = \mu + \text{مـ}_{\text{ل}} + \text{ق}_{\text{ر}} + \text{خ}_{\text{رل}} \qquad (٢/١)$$

حيث يعبر عن هذا النموذج فى صورة بيانات العينة على النحو التالى:

$$\bar{\text{س}}_{\text{رل}} = \bar{\text{س}}_{..} + (\bar{\text{س}}_{\text{ل.}} - \bar{\text{س}}_{..}) + (\bar{\text{س}}_{\text{.ر}} - \bar{\text{س}}_{..}) + \text{خ}_{\text{رل}}$$

$$= \hat{\mu} + \hat{\text{مـ}}_{\text{ل}} + \hat{\text{ق}}_{\text{ر}} + \text{خ}_{\text{رل}} \qquad (٢/٦)$$

حيث يعرف التأثير المضاف للمعالجات $\hat{\text{مـ}}_{\text{ل}} = (\bar{\text{س}}_{\text{ل.}} - \bar{\text{س}}_{..})$ (٢/٧)

والتأثير المضاف للقطاعات $\text{ق}_{\text{ر}} = (\bar{\text{س}}_{\text{.ر}} - \bar{\text{س}}_{..})$ (٢/٨)

بالفــرق بيــن متوسط المعالجة (ل) والوسط العام ، وكذلك متوسط القطاع (ر) والمتوسط العام. وحيث:

$$\sum (\bar{\text{س}}_{\text{ل}.} - \bar{\text{س}}_{..}) = \sum (\bar{\text{س}}_{.\text{ر}} - \bar{\text{س}}_{..}) = \text{صفر}$$

ومن (٢/١ ، ٦) يتبين الآتى:

$$\sum_{\text{ر ل}} \text{خ}^{2}_{\text{ر ل}} = \sum_{\text{رل}} (\text{س}_{\text{رل}} - \mu - \text{م}_{\text{ل}} - \text{ق}_{\text{ر}})^{2}$$

$$= \sum_{\text{ر ل}} (\text{س}_{\text{ر ل}} - \bar{\text{س}}_{\text{ل}.} - \bar{\text{س}}_{.\text{ر}} + \bar{\text{س}}_{..})^{2} \qquad (٢/٩)$$

وباستخدام أسلوب المربعات الصغرى يمكن تقدير قيم $\hat{\mu} = \bar{\text{س}}_{..}$ ،

$\hat{\text{م}}_{\text{ل}} = (\bar{\text{س}}_{\text{ل}.} - \bar{\text{س}}_{..})$ ، $\hat{\text{ق}}_{\text{ر}} = (\bar{\text{س}}_{.\text{ر}} - \bar{\text{س}}_{..})$

وهى تقديرات معالم النموذج (٢/١) بقيم وحيدة وهذه التقديرات هى:

$$\left. \begin{aligned} \hat{\mu} &= \frac{\text{م}}{\text{ر} \times \text{ل}} = \bar{\text{س}}_{..} \\ \text{م}_{\text{ل}} &= \frac{\text{م}_{\text{ل}}}{\text{ر}} - \bar{\text{س}}_{..} = \bar{\text{س}}_{\text{ل}.} - \bar{\text{س}}_{..} \\ \text{ق}_{\text{ر}} &= \frac{\text{م}_{\text{ر}}}{\text{ل}} - \bar{\text{س}}_{..} = \bar{\text{س}}_{\text{ر}.} - \bar{\text{س}}_{..} \end{aligned} \right\} \qquad (٢/١٠)$$

(١-٣) الكفاءة النسبية للنموذج:

كثــيراً مـــا يعنينا أن نتعرف على مدى كفاءة استخدام نموذج القطاعات الكاملــة العشــوائية ، كــبديل للنموذج العشوائى الكامل ، أى ما هو العائد من تخصــيص ل وحدة تجريبية عشوائية فى كل قطاع ، بدلا من توزيع ر ل وحدة عشــوائيا علـــى ل معالجـــة ، ويتمثل ذلك فى تخفيض خطأ التجربة أو الخطأ العشــوائى بطــريقة واضـــحة. ولتحقــيق ذلـــك يستخدم معامل الكفاءة النسبى (Efficiency) حيث:

م.ك. ن (قطاعات عشوائية / عشوائى كامل)= $\frac{\text{ع}_1^2}{\text{ع}_2^2}$ (١١/٢)

حيث ع_1^2: متوسط مربعات الخطأ العشوائى للنموذج كامل العشوائية.

ع_2^2: متوسط مربعات للخطأ العشوائى لنموذج القطاعات العشوائية.

ونظـــراً لأنـــه لا يتوفر لدينا فى هذه الحالة إلا ناتج التجربة لنموذج القطاعات العشوائية الكاملة ، لذلك تعين علينا تقدير ع_1^2 مـــن جدول تحليل التباين لهذا النموذج الأخير وهذا التقدير هو:

$$\hat{\text{ع}}_1^2 = \frac{\text{ر(ل} - ١) \text{ ع}_2^2 + \text{م . م . ق}}{\text{ر ل} - ١} \qquad (١٢/٢)$$

وبالتالى فإن :

$$\text{م.ك. ن.} = \frac{\hat{\text{ع}}_1^2}{\text{ع}_2^2} = \frac{\text{ر(ل} - ١) \text{ ع}_2^2 + \text{م . م . ق}}{\text{(ر ل} - ١) \text{ ع}_2^2} \qquad (١٣/٢)$$

وجدير بالذكر أن تقدير ع_1^2 من نتائج تحليل بيانات قطاعات عشوائية كاملة ≠ (مجموع مربعات القطاعات + مجموع مربعـــات البـــواقـــى) ÷ ل (ر−١) أى بإدمـــاج مجمـــوع مـــربعات القطاعات ودرجات حرية القطاعات إلى مجموع مـــربعات ودرجـــات حرية البواقى ، بفرض أن النموذج من البداية هو نموذج قطاعات عشوائية ، لأن ذلك إن صح ، فإن مجموع المربعات للمعالجات سوف يكون واحدا سواء كان النموذج المتبع فى تصميم التجربة ، هو نموذج العشوائية الكاملة الذى تخصص فيه ر ل وحدة تجريبية على جميع المعالجات عشوائياً أو نموذج القطاعات الكاملة العشوائية حيث توزع ل وحدة عشوائيا داخل كل قطاع، وذلك لأن مجموع المربعات الكلى واحد فى الحالتين واختلاف التصميم لابد وأن يؤدى إلى اختلاف المجموع للمعالجات.

وعـــادة مـــا يصحح معامل الكفاءة النسبى (١٣/٢) وذلك للتصحيح مقابل نقص درجـــات الحـــرية للخطأ العشوائى أوالبواقى باستخدام نموذج القطاعات الكاملة العشوائية وذلك على النحو التالى:

$$\text{م.ك.ن} = \frac{ع_1^2}{ع_2^2} \times \frac{(ن_2 + ١)\,(ن_1 + ٣)}{(ن_2 + ٣)\,(ن_1 + ١)} \times ١٠٠ \qquad (١٤/٢)$$

حيث

$ن_1$: درجات حرية البواقى لـــ $ع_1^2$ أى باستخدام النموذج العشوائى الكامل.

$ن_2$: درجات حرية البواقى لـــ $ع_2^2$ أى باســتخدام نموذج القطاعات الكاملة العشوائية.

ولقياس كفاءة استخدام القطاعات الكاملة العشوائية فى المثال الأخير فإن:

معامل الكفاءة النسبى (ع / ق ع)

$$= \frac{٥ \times ٢ \times ٦,٦٧ + ٤٠,٦٧}{١٤ \times ٦,٦٧} \times \frac{(٨+١)\,(١٢+٣)}{(٨+٣)\,(١٢+١)} \times ١٠٠$$

$$= \frac{٦٦,٧ + ٤٠,٦٧}{٩٣,٣٨} \times \frac{٩ \times ١٥}{١١ \times ١٣} \times ١٠٠$$

$$= ١,١٤٩٨ \times ٠,٩٤٤ \times ١٠٠ = ١٠٨,٥ \%$$

وهـــذا يعـــنى أن الخطـــأ العشوائى لو استخدم تصميم العشوائية الكاملة ووزعت الـــ ١٥ وحدة تجريبية عشوائياً على ثلاث معالجات ، لكان ١٠٨% مـــن قيمته لو استخدم أسلوب القطاعات العشوائية بتوزيع ثلاث وحدات تجريبية عشـــوائياً علـــى كل قطاع من القطاعات الخمسة ، وهذا يؤدى بالطبع إلى زيادة فـــترة الحـــد الأدنـــى للفـــروق المعنوية اللازمة لاختبار معنوية الفرق بين أى متوسطين ، كما أن ذلك يعنى أن استخدام ١٠٨ وحدة تجريبية باستخدام النموذج كامل العشوائية ، سوف يؤدى إلى الوصول إلى نفس الدرجة من الدقة باستخدام ١٠٠ وحـــدة فقـــط مع التحليل فى اتجاهين باتباع نموذج القطاعات العشوائية ،

وتزداد كفاءة القطاعات العشوائية كلما ازدادت معنوية أثر القطاعات كما سيتبين من المثال التالى.

مثال (٣-٣)*:

البـــيانات التالية تبين نتائج تجربة استخدم فيها نموذج القطاعات الكاملة العشوائية ، حيث اعتبرت أنواع الدهون (٨) معالجات ، اختبرت جميعها فى كل يـــوم مـــن الـــ (٦) أيام وهى تمثل القطاعات ، وكانت التجربة لقياس تباين أو اخـــتلاف نوع من الفطائر فى درجة امتصاصه للدهون بحسب النوع واليوم كما يلى:

	(نوع الدهون) معالجات								
الأيام (قطاعات)	١	٢	٣	٤	٥	٦	٧	٨	م ر
(١)	١٦٤	١٧٢	١٧٧	١٧٨	١٦٣	١٦٣	١٥٠	١٦٤.	١٣٣١
(٢)	١٧٧	١٩٧	١٨٤	١٩٦	١٧٧	١٩٣	١٧٩	١٦٩	١٤٧٢
(٣)	١٦٨	١٦٧	١٨٧	١٧٧	١٤٤	١٧٦	١٤٦	١٥٥	١٣٢٠
(٤)	١٥٦	١٦١	١٦٩	١٨١	١٦٥	١٧٢	١٤١	١٤٩	١٢٩٤
(٥)	١٧٢	١٨٠	١٧٩	١٨٤	١٦٦	١٧٦	١٦٩	١٧٠	١٣٩٦
(٦)	١٩٥	١٩٠	١٩٧	١٩١	١٧٨	١٧٨	١٨٣	١٦٧	١٤٧٩
م ل	١٠٣٢	١٠٦٧	١٠٩٣	١١٠٧	٩٩٠	١٠٥٨	٩٦٨	٩٧٤	٨٢٩٢ = م

م.م.ك = (١٦٤^{٢} + ١٧٢^{٢} + ... + ١٦٧^{٢}) − $\frac{٨٢٩٢^{٢}}{٤٨}$

= ١٤٤١٥٨٦ − ١٤٣٢٤٤٣ = ٩١٤٣,٠٠

* Anderson, R.L. and Bancroft T.A. Statistical Theory In Research, Mcgraw-Hill Book Company, 1952, pp. 238-239.

م.م. بين الدهون (معالجات) =

= $\frac{١}{٦}$ (١٠٣٢^{٢}+١٠٦٧^{٢}+٠٠٠+ ٩٧٤^{٢}) – ١٤٣٢٤٤٣

= ٣٣٤٤,٣٣

م.م. بين الأيام (القطاعات)=

= $\frac{١}{٨}$ (١٣٣١^{٢}+١٤٧٢^{٢}+٠٠٠+١٤٧٩^{٢}) – ١٤٣٢٤٤٣

= ٣٩٨٦,٧٥

م.م. المتبقى = ٩١٤٣ – ٣٣٤٤,٣٣ – ٣٩٨٦,٧٥ = ١٨١١,٩٢

جدول تحليل التباين

المصـــــدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
بين الأيـــــام (ق)	٣٩٨٦,٧٥	٥	٧٩٧,٣٥	١٥,٤٠
بين الدهون (م)	٣٣٤٤,٣٣	٧	٤٧٧,٧٦	٩,٢٣
البـــــــواقى	١٨١١,٩٢	٣٥	٥١,٧٧	ــــ
كلـــــــى	٩١٤٣,٠٠	٤٧	ــــ	ــــ

الفرض العدمى: $\bar{م}_{١}$ = $\bar{م}_{٢}$ = ... = $\bar{م}_{٨}$ = صفر

الفرض العدمى : $ق_{١}$ = $ق_{٢}$ = ... = $ق_{٦}$ = صفر

والفـــرض البديل فى كل حالة: هو عدم تساوى الأثر المتوسط لأى نوعين على الأقـــل مـــن الدهـــون (المعالجـــات) ، وكذلك للأثر المتوسط ليومين على الأقل (القطاعات).

ولكن $\text{ف}_{٥ ، ٣٥ ، ١\%}$ < ٣,٧٠ ، $\text{ف}_{٧ ، ٣٥ ، ١\%}$ < ٣,٣ لذلك يرفض الفرضين عند مستوى معنوية ١% .

فى هذا المثال سوف نجد أن:

معامل الكفاءة النسبى (ق ع / ع)

$$= \frac{٢.٤ \times ٥١,٧٧ + ٣٩٨٦,٧٥}{٤٧ \times ٥١,٧٧} \times \frac{٣٦ \times ٤٣}{٣٨ \times ٤١} \times ١٠٠ = ٢٥١,٥ \ \%$$

مما يعنى أن استخدام نموذج العشوائية الكاملة بتوزيع الدهون عشوائيا على الأيام الستة لكان الخطأ العشوائى (متوسط مربعات البواقى) قد ارتفع إلى ٢٥١ % من قيمته الحالية أو أننا كنا سنستخدم عددا من الوحدات التجريبية يصل إلى مرتين ونصف عدد الوحدات التى استخدمت فى هذه التجربة للوصول إلى نفس القيمة للخطأ العشوائي.

(١-٤) القراءات المفقودة Missing Plots:

قد تفقد إحدى القراءات أو أكثر نتيجة عوامل خارجة عن الإرادة ولا تتصل بظروف التجربة. وإذا كان النموذج المستخدم هو النموذج العشوائي الكامل ، فإن فقد قراءة أو أكثر لا يؤثر على التحليل من حيث الأسلوب متى كان عدد القراءات المتبقية لكل معالجة لا يقل عن ٢ ، إلا أن درجات الحرية للخطأ تنقص درجة مقابل كل وحدة تجريبية تفقد ، ويتم التحليل كالمعتاد سواء كان عدد الوحدات لكل معالجة متساو أو غير متساو . وجدير بالذكر أن تعدد القراءات المفقودة يستدعى مراجعة تخطيط التجربة وكيفية تنفيذها.

أما إذا كنا بصدد تجربة قطاعات كاملة عشوائية ، فإن فقد قراءة أو أكثر يعنى أن يفقد النموذج توازنه ، ولا تصبح القطاعات متعامدة فى أثرها مع المعالجات ، مما يستوجب تقدير القراءة (أو القراءات) المفقودة وإجراء التحليل

بعــد إحــلال القراءة المفقودة بتقديرها. وسوف نقتصر على كيفية معالجة تقدير قراءة مفقودة واحدة.

وأحد طرق تقدير القراءة المفقودة ، والتى سنرمز لها بالرمز س* والتى تجعل م.م. للبواقى أصغر ما يمكن هى:

$$س^{*} = \frac{ل \times م^{\bullet}_{ل} + ر \times م^{\bullet}_{ر} - م^{\bullet}}{(ر-1)\ (ل-1)} \qquad (٢/١٥)$$

حيث :

$م^{\bullet}_{ل}$: مجموع المعالجة المحتوية على القراءة المفقودة.

$م^{\bullet}_{ر}$: مجموع القطـــــاع المحتوى على القراءة المفقودة.

$م^{\bullet}$: المجموع الكلى:

ونضـــع س* مكــان القراءة المفقودة وتحلل نتائج التجربة على هذا الأساس مع تخفيض درجات الحرية للتباين الكلى والخطأ العشوائى كل بدرجة واحدة.

مثال (٢-٣):

نفــرض أن المفــردة الأخيرة فى العمود الأول من جدول مثال (٢-٣) وهى تمثل قيمة الوحدة التجريبية بعد المعالجة الأولى فى اليوم السادس قد فقدت (لــيس نتــيجة للتجربة وإلا كانت = صفر) . أوجد تقدير هذه القيمة واستكمل التحليل.

$$م^{\bullet}_{ر=٦} = ١٤٧٩ - ١٩٥ = ١٢٨٤$$

---

* لتقدير أكثر من قراءة مفقودة يرجع إلى بعض المراجع الأكثر تخصصا مثل:
Steel, R.G.D., Torrie; J.H.,; Pıinciples And Procedures of Statistics, McGraw – Hill Book company 1960.

م°ر-١ = ١٠٣٢ – ١٩٥ = ٨٣٧

م° = ٨٢٩٢ – ١٩٥ = ٨٠٩٧

إذن س* = (٨ × ٨٣٧ + ٦ × ١٢٨٤ – ٨٠٩٧) / (٥ × ٧)

= (٦٦٩٦ + ٧٧٠٤ – ٨٠٩٧) / ٣٥ = ١٨٠

أما مجاميع المركبات التى حسبت على أساس س* فهى كالآتى:

م.م.ك = ١٦٤^{٢} + ١٧٧^{٢} + ... + ١٨٠^{٢} + ١٧٢^{٢} + ... + ١٦٧^{٢}

– ٨٢٧٧^{٢} / ٤٨ = ١٤٣٥٩٦١ – ١٤٢٧٢٦٥ = ٨٦٩٦

م.م. بين الدهون (معالجات)

= ١/٦ (١٠١٧^{٢} + ١٠٦٧^{٢} + ... + ٩٧٤^{٢}) – ٨٢٧٧^{٢} / ٤٨

= ١٤٣٠٦٦٤,٨ – ١٤٢٧٢٦٥ = ٣٣٩٩,٨

م.م. بين الأيام (قطاعات)

= ١/٨ (١٣٣١^{٢}+١٤٧٢^{٢}+...+١٤٦٤^{٢}) – ٨٢٧٧^{٢} / ٤٨

= ١٤٣٠٩١١,٦ – ١٤٢٧٢٦٥ = ٣٦٤٦,٦

أذن م.م. البواقى = ١٦٤٩,٦

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
الأيام (قطاعات)	٣٦٤٦,٦	٥	٧٢٩,٣٢	١٥,٠٣
الدهون (معالجات)	٣٣٩٩,٨	٧	٤٨٥,٦٩	١٠,٠١
البواقى	١٦٤٩,٦	٣٤	٤٨,٥٢	
كلـى	٨٦٩٦,٠	٤٦	—	—

ولازال فرضى العدم مرفوضين عند α = ١% حيث:

$$\text{ف}_{\text{٥ ، ٣٤ ، ١\%}} < \text{٣,٧٠} \quad ، \quad \text{ف}_{\text{٧ ، ٣٤ ، ١\%}} < \text{٣,٣٠}$$

(١ – ٥) : الخطأ المعيارى للفرق بين أى متوسطين:

والخطأ المعيارى للفرق بين أى متوسطين هو كالمعتاد

$$\text{ع}_{\text{ف}} = \text{ع}_{\bar{\text{س}}_١ - \bar{\text{س}}_٢} = \sqrt{\text{ع}^٢ \left(\frac{١}{\text{ن}_١} + \frac{١}{\text{ن}_٢}\right)} = \sqrt{\text{ع}^٢ \left(\frac{٢}{\text{ن}}\right)} \qquad (١٦/٢)$$

حيث ن = ر إذا كان الخطأ المعيارى يحسب للفرق بين متوسط أى معالجتين أو ن = ل إذا كان يحسب للفرق بين متوسط أى قطاعين. (١ ، ٢) فى دليل $\bar{\text{س}}$ تشير إلى المعالجتين أو القطاعين موضوع المقارنة ، ع٢ هى متوسط مربعات البواقى فى جدول تحليل التباين.

وبالتالى فإن الحد الأدنى للفروق المعنوية لأى متوسطين *هو:

$$\left| \left\{ \text{ت}_{(\text{ر} - ١)(\text{ل} - ١) ، ١ - \alpha/٢} \right\} \text{ع} \left\{ \bar{\text{س}}_١ - \bar{\text{س}}_٢ \right\} \right| \qquad (١٧/٢)$$

وعليه فإن الخطأ المعيارى للفرق بين أى معالجتين من بيانات المثـــال (٢ – ٢) وجدول تحليل التباين له هو:

$$\text{ع}_{\bar{\text{س}}\,\text{ل}\,(\text{هـ})} - \text{ع}_{\bar{\text{س}}\,\text{ل}\,(\text{و})} = \sqrt{\text{٥١,٧٧} \times \frac{٢}{٦}} = \text{٤,١٥٤} \quad ، \quad \text{هـ} \neq \text{و}$$

والحد الأدنى للفروق المعنوية = |٣,٣٥ × ٤,١٥٤| = |١٣,٩٢|

وبالنسبة للقطاعات:

$$\text{ع}_{\bar{\text{س}}\,\text{ل}\,(\text{هـ})} - \text{ع}_{\bar{\text{س}}\,\text{ر}\,(\text{و})} = \sqrt{\text{٥١,٧٧} \times \frac{٢}{٨}} = \text{٣,٦٠} \quad ، \quad \text{هـ} \neq \text{و}$$

---

* حسبت هذه الفترة على أساس متوسط فترة الثقة للفرق بين أى متوسطين يختارا مقدما قبل إجراء التجربة وإلا كانت الفترة أكثر اتساعاً. ولمعالجة هذه النقطة يرجع إلى:

Anderson, R.L. Bancroft; T.A. "Statistical Theory in Research, McGraw – Hill Book Company 1952, P. 239.

والحد الأدنى للفروق المعنوية $= |٣,٣٥ \times ٣,٦٠| = |١٢,٠٦|$

وفى حالة وجود قراءة مفقودة فإن الخطأ المعيارى للفرق بين متوسط المعالجة ذات القراءة المفقودة ومتوسط أى معالجة أخرى يكون:

$$\text{ع}_{\text{ف}} = \sqrt{\text{ع}^{٢} \left(\frac{٢}{\text{ر}} + \frac{\text{ل}}{\text{ر}\,(\text{ر}-١)\,(\text{ل}-١)} \right)} \qquad (١٨/٢)$$

حيث ع ٢ : متوسط مربعات البواقى من جدول تحليل التباين المحسوب بعد تقدير القراءة المفقودة.

وتطبيقاً لذلك على المثال (٢ – ٣) فإن:

$$\text{ع}_{\text{ف}} = \sqrt{٤٨,٥٢ \left(\frac{٢}{٦} + \frac{٨}{٦ \times ٥ \times ٧} \right)}$$

$$= \sqrt{٤٨,٥٢ \,(٠,٣٣٣ + ٠,٠٣٨)} = ٤,٢٤$$

(٣) القطاعات الكاملة العشوائية مع التكرار:

(٣-١) مقدمة:

كثيراً ما تكون المشاهدة ليست هى وحدة التجربة ، ولكن الوحدة تشتمل علـــى عدة مفردات ، يسجل عن كل مشاهدة نتيجة إجراء التجربة، وحين تتعدد المشـــاهدات لكل وحدة تجربة أى يكون هناك عدة مشاهدات لكل قطاع / معالجة فإنه يتعين التمييز بين خطأ التجربة (Experimental Error) والخطأ العشوائى (Random error or Residual) وســـوف نقتصـــر على معالجة حالة تساوى عدد التكرارات (replicates) لكل وحدة تجربة ، أى لكل معالجة وقطاع. وفى هذه الحالة فإن النموذج الرياضى يكون كما يلى:

(٣-٣) النموذج الرياضى :

$$\text{س}_{\text{ر ل ك}} = \mu + \text{م}_{\text{ل}} + \text{ق}_{\text{ر}} + (\text{م ق})_{\text{ر ل}} + \text{خ}_{\text{ر ل ك}} \qquad (١٩/٢)$$

وهـذا النموذج لا يختلف فى فروضه عن سابقة (٢ -١) إلا فى إضافة الحـد (مـ ق) رر ، وهو يعبر عن اختلاف درجة الاستجابة للمعالجات المختلفة من قطاع لآخر. أو بعبارة أخرى فإنه يعنى وجود أثر لكل معالجة / قطاع (معا) على الوحدة التجريبية ، بالإضافة إلى الأثر الثابت لكل معالجة والأثر الثابت لكل قطاع وهو بهذا المفهوم يعرف باسم خطأ التجربة (Experimental Error) ، وهذا الأثر لا يعتبر عشوائياً بالطبع لأنه ينشأ من عوامل معروفة ذات أثر يمكن قياسـها وهـى اختلاف القطاعات وهى ليست عشوائية. وهذا ما يميز مجموع المربعات لخطأ التجربة عن مجموع المربعات الذى يعزى إلى الخطأ العشوائى أو خطأ العينة (Sampling Error) أو البواقى الذى يقيس اختلاف القراءات أو المشاهدات داخل أى وحدة تجربة عن بعضها البعض.

هـذا وقد أخذ فى الاعتبار أن النموذج المتبع فى تخطيط وتنفيذ التجربة هـو الـنموذج الثابت (Fixed Model) أى أن التجربة تشمل جميع المعالجات وجمـيع القطاعات الممكنتين وليست عينات عشوائية من مجتمعات أكبر وأوسع من المعالجات والقطاعات ، وبالتالى فإن الأثـر المضـاف لكل معالجة/ قطاع، يعتـبر فى حكم التفاعل الثـابت (Fixed interaction) . وطالما أن ك > ١ ، فإنه يمكن استخدام م.م.م. للبواقى أو الخطأ العشوائى لاختبار الفروض الخاصة بالمعالجـات والقطاعـات وخطـأ التجربة ، وهذا ما يميز النموذج الثابت عن النموذج العشوائى (Random Model) والذى تكون فيه المعالجات والقطاعات ليسـت سـوى عينة من مجتمع المعالجات والقطاعات ، وبالتالى فإن الاستنتاج الإحصـائى يقتصر على المعالجات والقطاعات التى أجريت عليها التجربة. أو الـنموذج المختلط (Mixed Model) حين تكون المعالجات أو القطاعات عينة مـن المجتمع الأكبر ويكون العنصر الآخر أى المعالجات أو القطاعات التى لم تعامل كعينات هى كافة ما هو متاح منها. وفى أى من هاتين الحالتين يكون مقام ف مختلف عن سابقة فى الحالة الأولى. ونكتفى بهذا القدر فى معالجة هذا الأمر.

ويكون جدول تحليل التباين فى حالة النموذج الثابت كالتالى:

(٢-٣) نموذج جدول تحليل التباين لتصميم القطاعات الكاملة العشوائية مع التكرار (النموذج الثابت) :

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية
القطاعات	$\sum_{\text{ر}} \frac{(\text{م}_{\text{ر..}})^{٢}}{\text{ل ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$	ر - ١
المعالجات	$\sum_{\text{ل}} \frac{(\text{م}_{\text{ل..}})^{٢}}{\text{ر ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$	ل - ١
خطأ التجربة	$\sum_{\text{رل}} \frac{(\text{م}_{\text{رل.}})^{٢}}{\text{ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$ - م.م.ق - م.م.م	(ر-١) (ل-١)
البواقى (خطأ عشوائى)	$\sum_{\text{ر ل ك}} \left(\sum_{\text{رل ك}} \text{س}^{٢} - \frac{(\text{م}_{\text{رل.}})^{٢}}{\text{ك}} \right)$	رل (ك -١)
الكلى	$\sum_{\text{ر ل ك}} \text{س}^{٢}_{\text{رل ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$	ر ل ك - ١

وسوف نستخدم بيانات المثال التالى (٢-٤) كشرح عملى لأسلوب التحليل فى هذه الحالة .

مثال (٢ -٤):

بفـــرض أن البيانات التالية تمثل إنتاج ثلاثة أنواع من الآلات (ل = ٣ معالجـــات) فـــى خمسة مواقع إنتاجية (ر = ٥ قطاعـــات) حيـــث سجلـــت قـــراءتين فـــــى كـــل موقع على كل آلة (أى أن الوحدة التجريبية تتكون من ك = ٢ مشاهدة).

أنواع الآلات (معالجات)	مواقع الإنتاج (القطاعات)					$م_ل$
	(١)	(٢)	(٣)	(٤)	(٥)	
أ	٢٥ ٢٢	٢٨ ٢٥	٢٩ ٢٠	٢٨ ٢٢	٢٦ ٢٠	٢٤٥
ب	٢٥ ٣٠	٢٤ ٣٠	٣٠ ٢٨	٣١ ٣٠	٢٧ ٢٥	٢٨٠
ج	٢٥ ٢٩	٣٠ ٢٠	٢٥ ٢٦	٢٣ ٢٨	٢٠ ٢٩	٢٥٥
$م_ر$	١٥٦	١٥٧	١٥٨	١٦٢	١٤٧	٧٨٠

مجموع المربعات الكلى = $٢٥^{٢} + ٢٨^{٢} + \cdots + ٢٩^{٢} - \frac{٧٨٠^{٢}}{٣٠}$

$= ٢٠٦٢٤ - ٢٠٢٨٠ = ٣٤٤$

مجموع المربعات بين القطاعات

$= \frac{١}{٣\times٢}(١٥٦^{٢}+١٥٧^{٢}+١٥٨^{٢}+١٦٢^{٢}+١٤٧^{٢}) - \frac{٧٨٠^{٢}}{٣٠}$

$= ٢٠,٣٣$

مجموع المربعات بين المعالجات

$= \frac{١}{٥\times٢}(٢٤٥^{٢}+٢٨٠^{٢}+٢٥٥^{٢}) - \frac{٧٨٠^{٢}}{٣٠} = ٦٥$

مجموع المربعات للتفاعلات (أو خطأ التجربة)

= $\frac{١}{٢}$ (٤٧^{٢}+٥٣^{٢}+...+٤٩^{٢}) − $\frac{٧٨٠^{٢}}{٢ \times ٣ \times ٥}$ − م.م.ق − م.م.مـ

= مجموع المربعات بين الخلايا (وحدات التجربة) − م.م.ق − م.م.مـ

= (٢٠٣٩٢ − ٢٠٢٨٠) − ٦٥ − ٢٠,٣٣ = ٢٦,٦٧

مجموع المربعات للخطأ العشوائى (أو البواقى)

= م.م. الكلى − م.م. بين الخلايا.

= ٣٤٤ − ١٢٢ = ٢٣٢

بدرجات حرية = ٢٩ − (٤+٢+٨) = ٢٩ − ١٤ = ١٥

ويكون جدول تحليل التباين كالآتى:

مصدر الاختلافات	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين ف*	ف α = ٥%
بين القطاعات	٢٠,٣٣	٤	٥,٠٨	٠,٣٣	٣,٠٦
بين المعالجات	٦٥	٢	٣٢,٥٠	٢,١٠	٣,٦٨
خطأ التجربة	٢٦,٦٧*	٨*	٣,٣٣	٠,٢٢	٢,٦٤
بين وحدات التجربة(الخلايا)	١١٢**	١٤**			
البواقى (الخطأ العشوائى)	٢٣٢	١٥		١٥,٤٧	
الكــلى	٣٤٤	٢٩	ـــــ	ـــــ	

* بالفرق ، ** بالفرق

وتقبل الفروض بأن مـ$_{١}$ = مـ$_{٢}$ = مـ$_{٣}$ = صفر

ق$_{١}$ = ق$_{٢}$ = ... = ق$_{٥}$ = صفر

(مـ ق)$_{١}$ = (مـ ق)$_{٢}$ = ... = ... = (مـ ق)$_{٩}$ = صفر

عند مستوى المعنوية ٥ %.

ثانياً : المربع اللاتينى
LATIN SQUARE

(١) مقدمة:

وفى هذا النموذج نوزع كل معالجة عشوائيا داخل كل صف وكل عمود، بحيــث تظهر كل المعالجات فى كل الصفوف وكل الأعمدة بشرط أن تظهر كل معالجـــة مرة واحدة وواحدة فقط داخل كل صف وكل عمود وبالتالى يمكن إزالة كـــل تباين نتيجة اختلاف الأعمدة والصفوف من الخطأ. وفى هذا النموذج يكون عدد الصفوف = عدد الأعمدة = عدد المعالجات = ر مثلاً. ففى تجربة تسويقية يمكـــن أن تكـــون الأيام هى الصفوف والمتاجر هى الأعمدة وأنواع السلع هى المعالجات كما أنه فى تجربة أخرى يمكن أن تقوم الأيام مقام القطاعات (٤ أيام أسبوعياً مثلا) وهى الأعمدة. كما أن الصفوف (٤) تمثل أنواع مختلفة من الآلات والورديـــات هـــى العامل الثالث (المعالجات وعددها ٤ أيضاً) ولنموذج المربع اللاتينى تطبيقات عديدة فى مجال التجارب الزراعية.

ويمكـــن أن نصـــل إلـــى عدد كبير من التوليفات الممكنة لكل عدد من المعالجـــات ، ففـــى حالة مربع لاتينى ٤ ×٤ فيوجد ٥٧٦ توليفة ممكنة أحداها النموذج التالى:

		المعالجات (الأعمدة)			
		I	II	III	IIII
قطاعات	(١)	أ	ب	جـ	د
	(٢)	د	أ	ب	جـ
	(٣)	جـ	د	أ	ب
	(٤)	ب	جـ	د	أ

حيـــث تمثل أ ، ب ، جـــ ، د الأنواع الأربعة للعامل الثالث الذى ادخل على الأعمدة والصفوف.

وفـــى حالـــة مربع لاتينى ٥ × ٥ **فأنه يوجد ١٦١٢٨٠ توليفة ممكنة ، يمكـــن اختـــيار** أحداها بالرجوع **إلى أحد المراجع المتخصصة الواردة فى قائمة المراجع فى نهاية هذا المرجع.**

والعـــدد المناسب لاستخدام **هذا النموذج يتراوح بين ٥×٥ إلى ٨×٨ إذا** أنه مع زيادة العدد عن ذلك ، سيصبح **عدد الوحدات التجريبية غير عملى. ومع** ذلـــك فإن عدد درجات حرية الخطأ لهذا النموذج = $ن^٢$ -١ أى أنه عند ن = ٣ فإن درجات حرية الخطأ = ٨ **ويعنى ذلك ضرورة مراعاة اختيار العدد المناسب** من الناحية العملية **ولتوفير العدد المناسب لدرجات حرية الخطأ.**

وسوف نستخدم بيانات المثال التالى (٢ – ٥) لعرض نتائج تجربة مربع لاتينى ٤ ×٤.

مثال (٣ –٥):

البيانات التالية تمثل نتائج تجربة مربع لاتينى ٤ × ٤:

		مواقع الإنتاج (قطاعات)				
		I	II	III	IIII	$م_{ص}$
أنواع الآلات (معالجات)	(١)	٨١/أ	٤١/ب	٤٤/جـ	٥٣/د	٢١٩
	(٢)	٣٨/د	٩٧/أ	٤٢/ب	٤٩/جـ	٢٢٦
	(٣)	٣١/جـ	٤٣/د	٦٧/أ	٣٦/ب	١٧٧
	(٤)	٥٧/ب	٣٣/جـ	٤٣/د	٨١/أ	٢١٤
$م_{ع}$		٢٠٧	٢١٤	٤٩٦	٢١٩	٨٣٦=م
الأيام		أ	ب	جـ	د	
$م_{ر}$		٣٢٦	١٧٦	١٥٧	١٧٧	٨٣٦=م

وتحليل النتائج يبين:

مجموع المربعات الكلى = $٨١^٢+٤١^٢+\cdots+٨١^٢ - \frac{٨٣٦}{٤×٤} = ٥٦٦٧$

مجموع المربعات للأعمـــدة (مواقع الإنتاج) = $\frac{١}{٤}(٢٠٧^{٢} + ٢١٤^{٢} + ١٨٦٠^{٢} + ٢١٩^{٢}) - \frac{٨٣٦}{٤\times٤} = ٧٤,٥$

مجمـــوع المربعـــات للصفـــوف (الآلات) = $\frac{١}{٤}(٢١٩^{٢} + ٢٢٦^{٢} + ١٧٧^{٢} + ٢١٤^{٢}) - \frac{٨٣٦}{٤\times٤} = ٣٥٩,٥$

مجموع المربعات لأيام = $\frac{١}{٤}(٣٢٦^{٢} + ١٧٦^{٢} + ١٥٧^{٢} + ١٧٧^{٢}) - \frac{٨٣٦}{٤\times٤} = ٤٦٢٦,٥$

ويكون مجموع المربعات المتبقى=٥٦٦٧ – ٧٤,٥ – ٣٥٩,٥ – ٤٦٢٦,٥

= ٦٠٦,٥

جدول تحليل التباين

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
الأعمدة (مواقع الإنتاج)	٧٤,٥	٣	٢٤,٨	٠,٢٥
الصفوف (لآلات)	٣٥٩,٥	٣	١١٩,٨	١,١٨
الأيـــام	٤٦٢٦,٥	٣	١٥٤٢,٢	١٥,٢٥
البـــواقى	٦٠٦,٥	٦	١٠١,١	—
الكـــلى	٥٦٦٧,٠	١٥	—	—

ولكـــن ف$_{٣ ، ٦ ، ٠٥\%}$ = ٤,٧٦ لذلك يقبل الفرض بعدم وجود أثر لكل من الأعمـــدة والصـــفوف ، ويرفض الفرض العدمى الخاص بالأيام بمعنى أن الأيام ذات أثر معنوى على وحدات التجربة.

هـــذا ويلاحظ أن هذا النموذج يختلف عن التحليل فى اتجاهين ، فى أنه يفترض عدم وجود تفاعل بين الصفوف والأعمدة ، كما أن عدد وحدات التجربة = ر × ر أى ر^{٢} رغم دخول ثلاثة عوامل فى التجربة مما كان يعنى استخدام ر × ر × ر وحدة تجريبية.

(٣) النموذج الرياضى والجبرى (الحسابى):

وعموما فإن نموذج المربع اللاتينى هو:

$$\text{س}_{\text{رلك}} = \mu + \text{م}_{\text{ل}} + \text{ق}_{\text{ر}} + \text{ع}_{\text{ك}} + \text{خ}_{\text{رلك}} \qquad (\text{٢٠/٢})$$

ر ، ل ، ك = ١ ، ٢ ، ...... ، ن

أمــا فــروض النموذج فهى نفس الفروض السابقة لعامل وعاملين ، مع إضافة العامل الثالث ع وهى نموذج الآثار المضافة ، أى أن:

$$\sum_{\text{رل}} (\text{س}_{\text{رلك}} - \bar{\text{س}}_{...})^{\text{٢}} = \text{ن} \sum (\bar{\text{س}}_{\text{ر}..} - \bar{\text{س}}_{...})^{\text{٢}}$$

$$+ \text{ن} \sum (\bar{\text{س}}_{.\text{ل}.} - \bar{\text{س}}_{...})^{\text{٢}} + \text{ن} \sum (\bar{\text{س}}_{..\text{ك}} - \bar{\text{س}}_{...})^{\text{٢}}$$

$$+ \sum (\text{س}_{\text{رلك}} - \bar{\text{س}}_{\text{ر}..} - \bar{\text{س}}_{.\text{ل}.} - \bar{\text{س}}_{..\text{ك}} + \text{٢}\,\bar{\text{س}}_{...})^{\text{٢}} \qquad (\text{٢١/٢})$$

أو

م.م.ك = م.م. بين الأعمدة + م.م. بين الصفوف + م.م. بين المعالجات

+ م.م. للخطأ العشوائى (أو البواقى) (٢٢/٢)

وتحسب مجاميع المربعات كالسابق.

كمــا أن الخطــأ مستقل عن المعالجات والقطاعات والعوامل ، وأن لها توزيـــع معتاد (صفر ، $\sigma^{\text{٢}}$) وبالتالى لم يحدث اختلاف فى أسلوب التحليل عما سبق وهو ما يتبين لنا من المثال (٢ – ٥).

هــذا وتتم العشوائية فى نموذج المربع اللاتينى باختيار توليفة من كافة التوليفات الممكنة بطريقة عشوائية*.

---

* يمكن الرجوع إلى جداول فيشر وتيتس للحصول على الصور الممكنة للمربع اللاتينى ٣×٣ حتى ٦×٦

Fisher, R.A. and Tates, F., Statistical Tables for Biological, Agricultural And Medical Research, 5 th ed., Hafner publishing company, N.Y., 1957.

(٣) الكفاءة النسبية لنموذج المربع اللاتينى:

لقياس كفاءة المربع اللاتينى مقارنا بالقطاعات الكاملة العشوائية يستخدم المعامل التالى:

$$\text{م.ك.ن(م لاتينى / ق ع)} = \frac{\text{ع}^{2*}}{\text{ع}^{2}} \times \text{معامل التصحيح} =$$

$$\frac{\text{م.م.ق} + (\text{ن}-1)^{2}\ \text{ع}^{2}}{\text{ن(ن}-1)\text{ع}^{2}} \times \frac{(\text{ن}_{2}+1)\ (\text{ن}_{1}+3)}{(\text{ن}_{2}+3)\ (\text{ن}_{1}+1)} \times 100 \qquad (23/2)$$

حيث:

ع^{2} : متوسط مربعات البواقى للمربع اللاتينى.

م.م.ق: مجموع مربعات القطاعات.

(ن −١) : درجات الحرية للقطاعات (أو المعالجات لانهما متساويان)

ن_{2} : درجات حرية ع^{2} للمربع اللاتينى

ن_{1} : درجات حرية ع^{2} للقطاعات الكاملة العشوائية.

وجدير بالذكر أننا نقدر كفاءة تصميم المربع اللاتينى من واقع ناتج تحليله مقارنا بالقطاعات العشوائية الكاملة تقديراً للخطأ العشوائى لو أننا استبدلنا من البداية نموذج المربع اللاتينى بالقطاعات الكاملة العشوائية وذلك يعتمد على مقارنة ع^{2*} للأخير مقدرة من ع٢ لتصميم المربع اللاتينى كما سبق أن بينا فى حالة القطاعات الكاملة العشوائية ، أما الحد الأخير فى (٢٣/٢) فهو تصحيح لنقص درجات الحرية باستخدام المربع اللاتينى.

وتطبيقاً لذلك على بيانات المثال (٢−٥) نجد الآتى:

$$\text{م.ك.ن} = \frac{47.5 + 9 \times 101.1}{4 \times 3 \times 101.1} \times \frac{(6+1)\ (9+3)}{(6+3)\ (9+1)} \times 100$$

$$\text{م.ك.ن} = \frac{47.5 + 909.9}{12 \times 101.1} \times \frac{7 \times 12}{9 \times 10} \times 100 = 75.7\ \%$$

أى أن القطاعـــات الكاملة العشوائية أعلى كفاءة من المربع اللاتينى فى حالتـــنا هذه (باستخدام الأعمدة كقطاعات) وأنه يمكن الوصول إلى نفس الدرجة مـــن الكفـــاءة (نفس الخطأ العشوائى) بـــ ٧٥ مفردة باستخدام القطاعات الكاملة العشوائية مقابل ١٠٠ مفردة باستخدام المربع اللاتينى.

وإذا اعتـــبرنا الصـــفوف هـــى القطاعـــات ، فأن كفاءة النموذج مقارنا بالقطاعات الكاملة العشوائية هى:

م.ك.ن (مربع لاتينى / ق / ع (للصفوف)

$$= \frac{\text{٣٥٩,٥} + \text{٩} \times \text{١٠١,١}}{\text{٤} \times \text{٣} \times \text{١٠١,١}} \times \frac{\text{٨٤}}{\text{٩٠}} \times \text{١٠٠} = \text{٩٧,٧}\%$$

وباعتبار الصفوف هى القطاعات ، فإن كفاءة المربع اللاتينى تقترب من كفـــاءة القطاعات الكاملة العشوائية. وهكذا تزداد كفاءة استخدام المربع اللاتينى كلما كانت القطاعات متباينة الأثر فيما بينها.

(٤) القراءة المفقودة:

وبفرض أن أحد القطع أو أحد الوحدات التجريبية فقدت لسبب آخر لا يرجع إلى الـــتجربة ذاتها ، فإنه يمكن تقدير قيمة هذه القراءة المفقودة بنفس الأسلوب الذى اتبع فى السابق ويكون تقديرها كالآتى:

$$\text{ص}^{*} = \frac{\text{ن}(\text{م}^{*}_{\text{١}} + \text{م}^{*}_{\text{٢}} + \text{م}^{*}_{\text{٣}}) - \text{٢}\,\text{م}^{*}}{(\text{ن} - \text{١})\,(\text{ن} - \text{٢})} \qquad (\text{٢٤/٢})$$

حيث $\text{م}^{*}_{\text{١}}$ ، $\text{م}^{*}_{\text{٢}}$ ، $\text{م}^{*}_{\text{٣}}$ ، م^{*} هى مجموع الصف والعمود والمعامل والكلى الذى يشتمل على القراءة المفقودة.

مثال (٣-٦):

وبفـــرض أن المفـــردة (I ، ٤ ، ب) أى المفردة فى نهاية العمود الأول وبداية الصف الرابع قد فقدت قيمتها الحقيقية فإن تقديرها يكون كالآتى:

$م^*_١ = ١٥٧$ ، $م^*_٢ = ١٥٠$ ، $م^*_٣ = ١١٩$ ، $م^* = ٧٧٩$

$$\text{أذن } ص^* = \frac{٤(١٥٧ + ١٥٠ + ١١٩) - ٢ \times ٧٧٩}{٣ \times ٢} = ٢٤,٣$$

وتستبدل القراءة الأخيرة المفقودة بالقيمة المقدرة ٢٤,٣ ويعاد التحليل من جديد، وتـــنقص درجات حرية المجموع الكلى والبواقى كل بدرجة واحدة مقابل تقدير القراءة المفقودة.

أمـــا الخطـــأ المعـــيارى للفرق بين متوسطى أى معالجتين تحتوى أحداهما على القراءة المفقودة فهو:

$$ع_{ف} = ع_{\bar{س}_و - \bar{س}_هـ} = \sqrt{ع^{٢*}\left(\frac{٢}{ن} + \frac{١}{(ن-١)(ن-٢)}\right)} \qquad (٢٥/٢)$$

و $\neq$ هـــ ، $ع^{٢*}$ = متوسط مربعات البواقى من جدول تحليل التباين.

ثالثاً: التحليل العاملى
Factorial Analysis

(١) مقدمة: العامل أو الأثر الأساسى و التفاعل:

التحليل العاملى (Factorial Analysis) هو تحليل نتائج تجارب تصمم خصيصاً بغرض إتاحة الفرصة لإجراء مقارنات مستقلة تعتمد على طريقة اختيار المعالجات ، والتى تمثلها فى هذا الأسلوب العوامل الرئيسية والتفاعلات ، وذلك باستخدام أسلوب تحليل التباين ، وكل عامل يتكون من عدة مستويات (Levels) ، ويعنينا أن نعين الأثر الأساسى المضاف لكل عامل عند مستوياته الممكنة ، وكذلك لكل مستوى للعامل مع المستويات المختلفة لعامل آخر ، ويعرف باسم التفاعل (Interaction) فمثلاً إذا اعتبر الحافز النقدى عاملاً له أثره على الإنتاج ، فأنه يمكن أن يتكون من عدة مستويات، وكل مستوى يحدد -مثلا- على أساس نسبة معينة من المرتب أو الأجر الشهرى، كما قد يعرف على أساس مبلغ معين من المال يصرف كحافز سنوى. وبالمثل إذا اعتبر البرنامج التدريبى عاملاً ، فأنه يمكن أن يتكون من عدة مستويات، فقد يكون البرنامج على مستوى المقدمة ثم برنامج متوسط المستوى وثالث متقدم.... وهكذا.

ولتصميم تجربة كهذه تشتمل على عامل واحد (Single Factor) مثل الحافز النقدى وأثره على الإنتاج ، وبفرض أن هذا العامل يتحدد بثلاثة مستويات هى ١٠% ، ٢٠% ، ٣٠% من الأجر الشهرى للعامل ، فأن ذلك يتطلب المحافظة على بقية المتغيرات – عدا الحافز النقدى – متجانسة أو ثابتة قدر الإمكان. وإذا ما أدخل عامل ثان فى الاعتبار كمدة الخبرة فى ممارسة العمل ، مثلاً ، وبفرض أنها قسمت إلى ثلاثة فترات (أو مستويات) هى أقل من ٥ سنوات ، من ٥ سنوات إلى أقل من ١٠ سنوات ، ثم ١٠ سنوات فأكثر، فأن ذلك

يعنى أننا بصدد تجربة مزدوجة العوامل، تتكون من ٣×٣ أو $٣^{٢}$ = تسعة صور ممكـــنة لمســـتويات العـــامل الأول مع مستويات العامل الثانى حيث يظهر كل مســـتوى من الحافز النقدى فى فترات الخبرة المختلفة ويمكن تصوير ناتج هذه التجربة بالجدول التالى:

ب: مدة الخبرة بالسنوات	أ : الحافز النقدى أ$_1$ ١٠%	أ$_2$ ٢٠%	أ$_3$ ٣٠%
ب$_1$: –٥	أ$_1$ ب$_1$	أ$_2$ ب$_1$	أ$_3$ ب$_1$
ب$_2$: ٥–	أ$_1$ ب$_2$	أ$_2$ ب$_2$	أ$_3$ ب$_2$
ب$_3$: ١٠ فأكثر	أ$_1$ ب$_3$	أ$_2$ ب$_3$	أ$_3$ ب$_3$

وقد استخدمت الحروف الأبجدية أ ، ب ، ... للدلالة على العوامل ويشير الدلـــيل أدنـــى كـــل حرف إلى المستوى الخاص بالعامل ، فقد استخدم الحرف أ للدلالة على العامل الأول وهو الحافز النقدى ، ب للعامل الثانى، وهو طول مدة الخبرة ، أما أ$_1$ فتشير إلى المستوى الأول أو الأدنى من العامل الأول أ ، أ$_2$ إلى المســـتوى الـــثانى لـــنفس العامل ... وكذلك ب$_1$ ترمز إلى أدنى مستوى للعامل الـــثانى ب ، وبالمـــثل ب$_2$ ترمـــز إلى المستوى الثانى لهذا العامل ، أما أ$_2$ ب$_2$ (مـــثلا) فتشـــير إلى أثر المستوى الثانى من العامل الأول أ مع المستوى الثانى للعـــامل ب أو المعالجة المشتركة للعاملين عند هذين المستويين، ويشار إليه فى التحليل العاملى بأنه تفاعل أ$_2$ مع ب$_2$ (Interaction) وهو يشير إلى التفاوت أو التبايـــن فى أثر أ عند المستويات المختلفة لـــ ب ، وهو يعنى أن هذا التباين لا يرجع إلى الصدفة أو العشوائية.

ويقوم أسلوب التحليل العاملى على أساس اعتبار المعالجات كعوامل ذات مســتويات مخــتلفة. وباســتخدام أســلوب تحليل التباين، يمكن تجزئة مجموع المــربعات للمعالجــات إلى مكونات بعضها يعبر عن الأثر الأساسى (Main effect) لكــل عــامل أ ، ب ، جـــ ، .... والبعض يعبر عن الأثر المشترك لعاملين أو أكثر أى التفاعلات ، ويمكننا ذلك من تحديد أفضل المستويات للعوامل فى إحداث أكبر أثر معنوى على وحدات التجربة.

وجدير بالذكر أنه حتى لو حدث وتبين عدم معنوية التفاعلات، فإن ذلك لا يقلــل من نتيجة التجربة أو إلى أهدار معلومات، بل أن ذلك يؤدى إلى تكرار أكبر للوحدات التجريبية لا يؤثر على دقة النتائج وإنما يعتبر تكرار مقنعا. وحين يتبين أن التفاعلات معنوية ، فأن ذلك يشير إلى عدم استقلال أثر العوامل.

وقــد عرفــت هذه النماذج فى البداية فى التجارب الزراعية ، ثم عرفت طــريقها بعــد ذلــك فى مجالات أخرى متعددة خاصة حين يكون الهدف تحديد المســتوى الأمــثل لعــامل ما مع مستويات عامل آخر ، كتحديد درجة الحرارة المناسبة مع الفترة الزمنية للوصول إلى درجة صلابة معينة للناتج أو أى تركيز مناسب لنوع من السماد مع أى نوع من البذور للوصول إلى أفضل محصول... وهكذا.

وسوف نقتصر فى معالجة النماذج التى سوف نقدمها عن التحليل العاملى فى الفقرات التالية ،على تخصيص عد متساو من الوحدات التجريبية لكل مستوى من العامل مع أى مستوى لعامل آخر مقتصرين فى ذلك على التجارب ٢×٢ أو $٢^{٢}$ أى لعاملين لكل مستويين.

(٣) التحليل العاملى لتجربة ٢×٢:

وفى هذه الحالة سوف يكون لدينا عاملين أ ، ب كلاهما له مستويين وبالتالى فأن المعالجــات ســوف تنقسم إلى الأثر الرئيسى لكل من أ ، ب ثم الأثر المشترك (الــتفاعل) لـــ أ ب وتعبر عنه المجموعات أ$_{١}$ ب$_{١}$ ، أ$_{١}$ ب$_{٢}$ ، أ$_{٢}$ ب$_{١}$ ، أ$_{٢}$ ب$_{٢}$

وسوف يكون النموذج الرياضى فى حالة استخدام أسلوب التحليل فى اتجاه واحد أو النموذج العشوائى الكامل كالآتى:

(٣-١) النموذج الرياضى:

س ر ك ر = μ + أ ر + ب ك + (أب) ر ك + خ ر ل ك (٢٦/٢)

حيث ترمز أ ر إلى الأثر المضاف للعامل أ ر ، (ل = ١ ، ٢) ، ب ك إلى الأثر المضاف للعامل الثانى ب ك (ك = ١ ، ٢)، ر فترمز إلى عدد وحدات التجربة لكل خلية (أ ر ، ب ك) ، أما خ ر ل ك فتشير إلى التباين العشوائى ، وسوف نفترض أن النموذج لآثار مضافة وأن خ ر ل ك مستقلة عن أ ، ب وأن لها توزيع معتاد توقعه الصفر وتباينه σ^2 كالمعتاد . كما أن (أ ب) ر ك فتشير إلى تفاعل أ مع ب أو اختلاف أثر أ مع مستويات ب ، وأن :

مجـ (أ ر) = مجـ (ب ك) = مجـ ر (أ ب) ر ك = مجـ ك (أ ب) ر ك = صفر.

وسوف نقتصر فى معالجة حساب مجموع المربعات للمعالجات وتجزئته إلى مجموع المربعات للعوامل وللتفاعلات بنفس الأسلوب الذى اتبع فى السابق وأن كانت هناك طريقة أخرى تعتمد على ما يعرف باسم المعاملات الخطية المتعامدة أو كثيرات الحدود المتعامدة (Orthogonal Polynomials or Orthogonal Linear Forms) ولكننا سوف لا نتعرض لهذه الطريقة الأخيرة فى الوقت الحاضر.

مثال (٣-٧)*:

لدراسة أثر العامل (أ) على تركيز مادة ما عند مستويات مختلفة من العامل (ب) وكان كل من العاملين يتميز بمستويين فقط، فقد خصص ر = ٥ وحدات تجريبية عشوائيا للتجميعات الأربعة للعاملين أ١ ب١ ، أ١ ب٢ ، أ٢ ب١ ، أ٢ ب٢ وكانت النتائج كالآتى:

* Steel m R.G.D., Torrie, J.H., : Principles and Procedures of Statistics, McGraw-Hill Book Company Inc., 1960, PP 199 – 202.

	أ$_{١}$		أ$_{٢}$		
	أ$_{١}$ ب$_{١}$	أ$_{١}$ ب$_{٢}$	أ$_{٢}$ ب$_{١}$	أ$_{٢}$ ب$_{٢}$	
	٨,٥٣	١٧,٥٣	٣٩,١٤	٣٢,٠٠	
	٢٠,٥٣	٢١,٠٧	٢٦,٢٠	٢٣,٨٠	
	١٢,٥٣	٢٠,٨٠	٣١,٣٣	٢٨,٨٧	
	١٤,٠٠	١٧,٣٣	٤٥,٨٠	٢٥,٠٦	
	١٠,٨٠	٢٠,٠٧	٤٠,٢٠	٢٩,٣٣	
مجـ س	٦٦,٣٩	٩٦,٨٠	١٨٢,٦٧	١٣٩,٠٦	٤٨٤,٩٢
مجـ س^{٢}	٩٦٣,٨٨	١٨٨٧,٠٢	٦٩١٣,٦٣	٣٩١٢,١٢	١٣٦٧٦,٧٠

ولتحليل نتائج هذه التجربة فأن ذلك يمكن أن يتم على خطوتين:

الخطوة الأولى :

وتحسب فيها مجموع المربعات المعتاد لتحليل التباين وهى:

مجموع المربعات الكلى = م. م. ك = مجـ س^{٢} $_{\text{ر ل ك}}$ − $\frac{(\text{م})^{٢}}{\text{ر ل ك}}$

= ١٣٦٧٦,٧٠ − $\frac{(٤٨٤,٩٢)^{٢}}{٥ \times ٢ \times ٢}$ = ١٩١٩,٣٣

مجموع المربعات للمعالجات* =

$\frac{١}{٥}$ (٦٦,٣٩^{٢} + ٩٦,٨٠^{٢} + ١٨٢,٦٧^{٢} + ١٣٩,٠٦^{٢})

− $\frac{(٤٨٤,٩٢)^{٢}}{٥ \times ٢ \times ٢}$ = ١٥٣٩,٤١

مجموع المربعات المتبقى = ١٩١٩,٣٣ − ١٥٣٩,٤١ = ٣٧٩,٩٢.

---

*ونحــن هنا بصدد ٤ معالجات يرمز لها بــ أ$_{١}$ ب$_{١}$ ، أ$_{١}$ ب$_{٢}$ ، أ$_{٢}$ ب$_{١}$ ، أ$_{٢}$ ب$_{٢}$ خصص لكل عــدد ٥ وحدات تجريبية. وبالتالى فأن م.م. للمعالجات يتكون من ثلاثة مركبات هى م.م.أ ، م.م.ب ، م م (أ×ب).

الخطوة الثانية:

وفيها يتم تجزئة مجموع المربعات للمعالجات إلى مجموع مربعات لكل عامل أساسى وأخر لكل من التفاعلات كالآتى:

مجموع المربعات للعامل (أ) أى م.م. (أ) $= \frac{\text{مجـ}_{\text{ل(أ)}} (\text{م}_{\text{ل (أ)}})^{٢}}{\text{ر} \times \text{ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$

واختصاراً للرموز $= \frac{\text{مجـ}_{\text{ل}} (\text{م}_{\text{أ}})^{٢}}{\text{ر} \times \text{ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$ (٢/٢٧)

حيث** : $\text{أ}_{\text{ر}} = ٢ = \text{أ}_{١} ، \text{أ}_{٢}$

$$\text{م.م.(أ)} = \frac{(٦٦,٣٩ + ٩٦,٨٠)^{٢} + (١٨٢,٦٧ + ١٣٩,٠٦)^{٢}}{٢ \times ٥} - \frac{(٤٨٤,٩٢)^{٢}}{٥ \times ٢ \times ٢}$$

$$= \frac{١٦٣,١٩^{٢} + ٣٢١,٧٣^{٢}}{١٠} - ١١٧٥٧,٣٧ = ١٢٥٦,٧٥$$

مجموع المربعات للعامل (ب) أى م.م. (ب) $= \frac{\text{مجـ}_{\text{ل(ب)}} (\text{م}_{\text{ل(ب)}})^{٢}}{\text{ر} \times \text{ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$

واختصاراً للرموز $= \frac{\text{مجـ}_{\text{ل}} (\text{م}_{\text{ب}})^{٢}}{\text{ر} \times \text{ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}}$ (٢/٢٨)

* حيث أن $\text{ب}_{\text{ر}} = \text{ب}_{١} ، \text{ب}_{٢}$

$$\text{م.م.(ب)} = \frac{(٦٦,٣٩ + ١٨٢,٦٧)^{٢} + (٩٦,٨٠ + ١٣٩,٠٦)^{٢}}{٢ \times ٥} - \frac{(٤٨٤,٩٢)^{٢}}{٥ \times ٢ \times ٢}$$

** رمــزنا إلى مستويات كل من العامل (أ) والعامل (ب) وعدد كل ٢ بالرمز ل ولو اختلف عــدد مســتويات أ عــن ب لاضطررنا لاستخدام رمز آخر مثل ل للدلالة على مستويات أ ، و (مثلا) للدلالة على مستويات ب.

$$= \frac{\text{٢٤٩,٠٦}^{٢} + \text{٢٣٥,٨٦}^{٢}}{\text{١٠}} - \text{١١٧٥٧,٣٧} = \text{٨,٧١}$$

مجموع المربعات للتفاعلات (أ ب) أى م.م. (أ ب) =

$$= \frac{\text{مجـ}_{\text{ل(أب)}} (\text{م}_{\text{أب}})^{٢}}{\text{ك}} - \frac{(\text{م})^{٢}}{\text{ر} \times \text{ل} \times \text{ك}} - \text{م . م (أ) – م.م (ب)} \qquad (\text{٢٩/٢})$$

$$= \frac{\text{٦٦,٣٩}^{٢} + \text{٩٦,٨٠}^{٢} + \text{١٨٢,٦٧}^{٢} + \text{١٣٩,٠٦}^{٢}}{\text{٥}} - \frac{(\text{٤٨٤,٩٢})^{٢}}{\text{٥} \times \text{٢} \times \text{٢}} - \text{١٢٥٦,٧٥} - \text{٨,٧١}$$

= مجموع المربعات للمعالجات – م . م (أ) – م. م. (ب) (٣٠/٢)

= ١٥٣٩,٤١ – ١٢٥٦,٧٥ – ٨,٧١ = ٢٧٣,٩٥

ثم يتم تصوير النتائج فى جدول تحليل التباين كالأتى:

المصدر للتباين	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
العامل (أ)	١٢٥٦,٧٥	ر–١ = ١	١٢٥٦,٧٥	٥٢,٩٢
العامل (ب)	٨.٧١	ل–١ = ١	٨,٧١	٠,٣٧
التفاعل (أب)	٢٧٣,٩٥	(ر–١) (ل–١) = ١	٢٧٣,٩٥	١١,٥٣
المعالجات	١٥٣٩,٤١	(رل–١) =٣	—	—
البواقى	*٣٧٩,٩٢	رل (ك–١) = ١٦*	٢٣,٧٥	—
الكلى	١٩١٩,٣٣	رل ك–١ = ١٩	–	–

* أو بالفرق

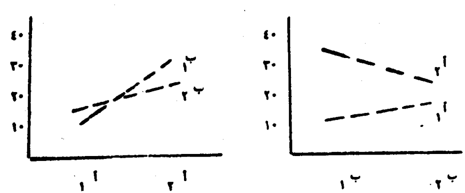
ويتبين من التحليل أن العامل (أ) والتفاعلات (أب) كلاهما له أثره المعنوى عند مستوى المعنوية ١% ، حيث ($\text{ف}_{\text{١ ، ١٦ ، ١\%}} = \text{٨,٥٣}$).

ومعنوية أثر التفاعلات تعنى أن العاملين أ ، ب غير مستقلى الأثر ، أى أن الفرق بين الأثر الأساسى للعامل أ عند المستويين للعامل ب أى عند ب$_1$، ب$_2$ معـــنوى ، والعكس أيضاً صحيح بمعنى أن الفرق بين الأثر الأساسى للعامل ب عند المستويين أ$_1$ ، أ$_2$ معنوى.

وإذا رجعنا إلى متوسطات المعالجات لوجدنا أن:

	أ$_1$ ب$_1$	أ$_1$ ب$_2$	أ$_2$ ب$_1$	أ$_2$ ب$_2$
$\bar{س}$	١٣,٢٨	١٩,٣٦	٣٦,٥٣	٢٧,٨١

أى أن متوسط الأثر للعامل أ عند المستوى ب$_1$ قد ارتفع من ١٣,٢٨ إلى ٣٦,٥٣ بالانتقال من المستوى الأدنى لـــ أ إلى المستوى الأعلى لذات العامل، أما عـــند المســـتوى ب$_2$ فقد ارتفع المتوسط من ١٩,٣٦ إلى ٢٧,٨١ . فى حين أنه بالنسبة للأثر المتوسط للعامل ب فقد انتقل من ١٣,٢٨ إلى ١٩,٣٦ عند المستوى الأدنـــى للعـــامل أ أى أ$_1$ ، أما عند المستوى الأعلى لـــ أ أى أ$_2$ فقد انتقل الأثر المتوسط للعامل ب من ٣٦,٥٣ إلى ٢٧,٨١ وهو يبين عدم استقلال أثر كل من العامليـــن عن العامل الآخر بل ويختلف الأثر اتجاها وحجما. ويمكن التعبير عن استجابة كل من العاملين عند كل مستوى للعامل الآخر بالشكل البيانى التالى:



ويمكـــن مـــن دراســـة هذا الشكل البيانى او من استكمال تحليل التباين للوصول إلى مجموع مربعات العامل (أ) مع أو من خلال (ب$_1$) وكذلك مجموع مـــربعات (أ) من خلال (ب$_2$) ، ومجموع مربعات العامل (ب) من خلال (أ$_1$)، وأخيراً مجموع مربعات العامل (ب) من خلال (أ$_2$) وهى:

م.م. (أ من خلال ب$_{1}$) = م.م. (أ/ب$_{1}$) = $\frac{\text{(م أ}_{2}\text{ / ب}_{1}\text{ − م أ}_{1}\text{ / ب}_{1}\text{)}^{2}}{\text{ل × ك}}$

= $\frac{(٦٦,٣٩ - ١٨٢,٦٧)^{2}}{٢ \times ٥}$ = ١٣٥٢,١٠ (٣١/٢)

وكذلك م.م. (أ / ب$_{2}$) = $\frac{\text{(م أ}_{2}\text{ | ب}_{2}\text{ − م أ}_{1}\text{ | ب}_{2}\text{)}^{2}}{\text{ل × ك}}$ (٣٢/٢)

= $\frac{(٩٦,٨٠ - ١٣٩,٠٦)^{2}}{٢ \times ٥}$ = ١٧٨,٥٩

أمـــا م.م. (ب / أ$_{1}$) = $\frac{\text{(م ب}_{2}\text{ | أ}_{1}\text{ − م ب}_{1}\text{ | أ}_{1}\text{)}^{2}}{\text{ر × ك}}$ (٣٣/٢)

= $\frac{(٦٦,٣٩ - ٩٦,٨٠)^{2}}{١٠}$ = ٩٢,٤٨

وكذلك م.م. (ب / أ$_{2}$) = $\frac{\text{(م ب}_{2}\text{ | أ}_{2}\text{ − م ب}_{1}\text{ | أ}_{2}\text{)}^{2}}{\text{ر × ك}}$ (٣٤/٢)

= $\frac{(١٨٢,٦٧ - ١٣٩,٠٦)^{2}}{١٠}$ = ١٩٠,١٨

ويمكـــن مـــن خلال ذلك الوصول إلى قرار بشأن أفضل جمع لمستوى للعامل (أ) مع مستوى للعامل (ب). وبتحليل التباين لهذه النواتج يتبن الآتى:

مقارنة المعالجات	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
أ من خلال أو مع ب$_{1}$	١٣٥٢.١٠	ر −١ = ١	١٣٥٢,١٠	٥٦,٩٣
أ مع ب٢	١٧٨,٥٩	ر −١ =١	١٧٨,٥٩	٧,٥٢
ب مع أ$_{1}$	٩٢,٤٨	ل−١=١	٩٢,٤٩	٣,٨٧
ب مع أ$_{2}$	١٩٠,١٨	ل−١=١	١٩٠,١٨	٨,٠١
البواقى	٣٧٩,٩٢	١٦	٢٣,٧٥	

ولكن ف$_{١ ، ١٦ ، ١\%}$ = ٨,٥٣ ، ف$_{١ ، ١٦ ، ٥\%}$ = ٤,٤٩

ممـــا يمكـــن معــــه القول بأن أعلى المعالجات أثرا هى للعامل (أ) عند المســـتوى الأدنــــى للعـــامل (ب) ، أما أثر (أ) مع المستوى الأعلى لــ (ب) أو للعـــامل (ب) مع المستوى الأعلى للعامل (أ) فهو أقل معنوية. أما بالنسبة للعامل (ب) عند المستوى الأدنى للعامل (أ) فهو غير معنوى.

وجديـــر بالإشارة إلى أن مجموع المربعات للعامل عند مستويات العامل الآخر = مجموع المربعات للأثر البسيط أو الأساسى للعامل مضافا إليه مجموع المربعات الذى ينشأ من تفاعل هذا العامل مع العامل الآخر فمثلا:

م.م. (أ/ب$_{١}$) + م.م. (أ/ب$_{٢}$) = ١٣٥٢,١٠ + ١٧٨,٥٩ = ١٥٣٠,٦٩

= م.م. (أ) + م.م. (أب) = ١٢٥٦,٧٥ + ٢٧٣,٩٥ =١٥٣٠,٦٩

وبالمثل فأن:

م.م. (ب/أ$_{١}$) + م.م. (ب/أ$_{٢}$) = ٩٢,٤٨٠ + ١٩٠,١٨ = ٢٨٢,٦٦

= م.م. (ب) + م.م. (أ ب) = ٨,٧١ + ٢٧٣,٩٥ = ٢٨٢,٦٦

تمـارين

(١) البـيانات التالية تبين عدد الوحدات المعيبة فى كل ١٠٠ وحدة أنتجت فى خمسة أيام متتالية فى أربعة مواقع إنتاجية مختلفة:

مواقع الإنتاج	الأيـــــام				
	(١)	(٢)	(٣)	(٤)	(٥)
أ	٧	٨	٨	١٠	١٢
ب	٥	٥	٦	٦	٨
جـ	٦	٧	٨	٩	١٠
د	٥	٧	٧	٨	٨

حلـل نـتائج هـذه الـتجربة باستخدام أسلوب تحليل التباين عند مستوى المعنوية ٥% . حدد النموذج والفروض التى تختبرها ثم أوجد الحد الأدنى للفروق المعنوية لكل من الأيام ومواقع الإنتاج عند α = ٥%.

(٢) لدراســة أثر أسلوب الإعلان وطريقة التعليب على الطلب على سلعة ما فقد سجلت هذه النتائج من تجربة عشوائية على السلعة فى ستة منافذ للتوزيع لأحد المتاجر:

	طريقة التعليب		
أسلوب الإعلان	أ	ب	جـ
صحافة	٣٠	٣٧	٣٢
تليفزيون	٣٤	٣٨	٤٢

تحقـق من معنوية أثر أسلوب الإعلان وطريقة التعليب على الطلب على السلعة باستخدام تحليل التباين وذلك عند مستوى المعنوية ٥%.

كيف يقارن هذا النموذج بالنموذج كامل العشوائية من حيث الكفاءة ؟.

(٣) وبفــرض أن تجــربة السؤال السابق قد كررت فى عدة أيام متتالية فكانت النتائج كالآتى:

	طريقة التعليب		
أسلوب الإعلان	أ	ب	جـ
صحافة	٣٠	٣٧	٣٢
	٣١	٣٥	٣٠
	٢٨	٣٩	٢٧
تليفزيون	٣٤	٣٨	٤٢
	٣٨	٣٧	٤١
	٣٦	٤٠	٤٦

عيــن الــنموذج الخــاص بهذه التجربة، ثم حلل نتائجها مستخدماً تحليل التباين. بماذا توصى فى ضوء نتائج التحليل؟

(٤) لــتقدير ومقارنة ارتفاع الكفاءة لبرنامج تدريبى بحسب طول فترة التدريب، فقــد طــبق الــبرنامج التدريبى على أربعة مجموعات متساوية العدد من المتدربين فى ثلاثة أشهر متتابعة، وسجل الوقت المنفق فى إنتاج سلعة ما بالساعة فكانت النتائج كالآتى:

الشهر	المجموعة (١)	(٢)	(٣)	(٤)
أبريل	١٦	١٣	١٥	٩
مايو	٣٠	٢٢	٢٩	٢٣
يونيو	٢٤	٢٠	١٨	٢٤

حلل هذه البيانات مستخدما تحليل التباين، ثم أوجد الخطأ المعيارى للفروق بين متوسطات الأشهر. قارن بين كفاءة نموذج القطاعات الكاملة العشوائية المستخدم فى تصميم هذه التجربة والنموذج الكامل العشوائية.

(٥) وبفــرض أن القراءة الخاصة بالمجموعة الرابعة فى شهر يونيو فى السؤال السابق قد فقد تسجيلها، قدر هذه القيمة وأعد التحليل مرة أخرى عند نفس مستوى المعنوية.

(٦) البــيانات التالــية تبيــن الوقت (مقربا إلى أقرب ساعة) الذى أستغرق فى الإجابــة علــى اختبارين مختلفين طبقا فى ستة أيام متتالية ، وقد أجرى الاختبار مرتين فى كل حالة وكانت النتائج كالآتى:

الاختبار	اليوم (١)	(٢)	(٣)	(٤)	(٥)	(٦)	المجموع
أ	٧	٩	٨	٦	٨	٧	٩٦
	٩	١٠	٨	٨	٩	٧	
ب	٨	١٠	٧	١٠	٧	٨	١٠٥
	٩	١١	٨	١٠	٨	٩	
المجموع	٣٣	٤٠	٣١	٣٤	٣٢	٣١	٢٠١

حلــل نــتائج هذه التجربة وناقش نتائج التحليل، وبماذا توصى فى ضوء نتائج التحليل؟

(٧) أراد أحد معارض السيارات أن يقارن بين خمسة أنواع من السيارات حسب متوســط عــدد الكيلومترات / لتر بنزين فأختار عينة عشوائية مكونة من ثــلاث سيارات من كل نوع من ثلاث مدن كبيرة القاهرة / الإسكندرية / المنصـــورة وأجــرى الاختبار على أساس لتر واحد من البنزين فى كل سيارة من كل نوع فى كل مدينة وكانت النتائج كالآتى:

المدينة	نوع السيارة				
	أ	ب	جـ	د	هـ
القاهرة	٢٠,٣	١٩,٥	٢٢,١	١٧,٦	٢٣,٦
	١٩,٨	١٨,٦	٢٣,٠	١٨,٣	٢٤,٥
	٢١,٤	١٨,٩	٢٢,٤	١٨,٢	٢٥,١
الإسكندرية	٢١,٦	٢٠,١	٢٠,١	١٩,٥	١٧,٦
	٢٢,٤	١٩,٩	٢١,٠	١٩,٢	١٨,٣
	٢١,٣	٢٠,٥	١٩,٨	٢٠,٣	١٨,١
المنصورة	١٩,٨	١٩,٦	٢٢,٣	١٩,٤	٢٢,١
	١٨,٦	١٨,٣	٢٢,٠	١٨,٥	٢٤,٣
	٢١,٠	١٩,٨	٢١,٦	١٩,١	٢٣,٨

حلل نتائج هذه الدراسة موضحاً النموذج وفروضه.

(٨) اســتخدم نموذج المربع اللاتينى ٤ ×٤ لمقارنة أربعة طرق مختلفة لإنتاج سلعة ما، وسجلت عدد الوحدات المنتجة وكانت النتائج كالاتى:

د ١٢	أ ١١	جـ ١٣	ب ١٢
ب ١٤	جـ ١٣	أ ١٢	د ١٥
أ ١٠	ب ٩	د ١٠	جـ ١١
جـ ١٢	د ١٣	ب ١١	أ ١٢

بيــن أن اخــتلاف الطرق ليس له أثر معنوى على لإنتاج. ثم قارن بين الكفاءة النسبية لهذا النموذج ونموذج القطاعات الكاملة العشوائية فى ضوء نتائج التجربة وبماذا توصى أذن؟

(٩) وبفــرض أن المفــردة الــتى تقع فى الصف الرابع العمود الثانى قد سقط تسجيلها ، قدر هذه المفردة ثم أعد التحليل ثم أوجد الخطأ المعيارى للفرق بين متوسط الطريقة (د) وأى طريقة أخرى.

(١٠) الآتى بيانات مربع لاتينى ٦×٦ لمقارنة ستة أنواع من البذور:

							مجموع الصفوف
	ب ٢٢٠	و ٩٨	د ١٤٩	أ ٩٢	هـ ٢٨٢	جـ ١٦٩	١٠١٠
	أ ٧٤	هـ ٢٣٨	ب ١٥٨	جـ ٢٢٨	و ٤٨	د ١٨٨	٩٣٤
	د ١١٨	جـ ٢٧٩	و ١١٨	هـ ٢٧٨	ب ١٧٦	أ ٦٥	١٠٣٤
	هـ ٢٩٥	ب ٢٢٢	أ ٥٤	د ١٠٤	جـ ٢١٣	و ١٦٣	١٠٥١
	جـ ١٨٧	د ٩٠	هـ ٢٤٢	و ٩٦	أ ٦٦	ب ١٢٢	٨٠٣
	و ٩٠	أ ١٢٤	جـ ١٩٥	ب ١٠٩	د ٧٩	هـ ٢١١	٨٠٨
مجموع الأعمدة	٩٨٤	١٠٥١	٩١٦	٩٠٧	٨٦٤	٩١٨	٥٦٤٠

حلل نتائج هذه التجربة. أى أنواع البذور أكبر أثرا على إنتاجية المحصول ثـم قـارن بيـن كفاءة هذه التجربة وكفاءتها لو أنها صممت على أساس قطاعات عشوائية كاملة.

(١١) أجريت تجربة لاختبار تباين أثر أربعة أنواع من السماد أ ، ب ، جـ ، د وذلـك بترتيـب النـبات الذى يسمد فى شكل مربع لاتينى ٤×٤ أى أن الصـفوف والأعمدة كانت تمثل اتجاهات مختلفة فى حقل التجربة وكانت نتائج التجربة :

أ ١٧	ب ١٣	جـ ١٩	د ١٦
ب ١٣	جـ ٢٠	د ١٤	أ ١٨
جـ ٢١	د ١٦	أ ١٥	ب ١٢
د ١٤	أ ١٦	ب ١٤	جـ ١٨

حلــل هــذه البــيانات للتحقق من أن متوسط الغلة متباين باختلاف نوع السـماد وذلــك عند مستوى المعنوية ٥% . ثم قارن بين كفاءة النموذج مقارنا بنموذج القطاعات العشوائية الكاملة.

(١٢) يوضح الجدول التالى التحليل الجزئى للتباين لتجربة عاملية ذات عاملين أ،ب:

المصدر	م.م.	د.ح.	م.م.م.
العامل أ	...	٣	٠,٧٥
العامل ب	٠,٩٥	١	...
التفاعل (أ ×ب)	...	...	٠,٣٠
الخطأ	...	...	...
كلى	٦,٥	٢٣	

والمطلوب:

أولاً : تحديد عدد مستويات كل من العاملين وكذلك عدد وحدات التجربة عند كل مستوى لعامل مع مستويات العامل الآخر.

ثانــياً : ماذا يعنى " تفاعل عامل مع آخر" وكيف يمكن أن يستفاد منه عند تقييم نتيجة التجربة ؟

ثالــثاً: اســتكمل جــدول تحليل التباين السابق ثم اختبر معنوية الأثر المتوسط للمعالجات عند مستوى المعنوية α = ٠,٠١.

رابعــاً : اختــبر معنوية أثر كلاً من العاملين وكذلك تفاعلهما عند ذات مستوى المعــنوية ١% . وأى مــن الاختبارات السابقة (فى ثالثاً أو رابعاً) تحققها التحليل العاملى ؟

الفصل الثالث

الانحدار الخطي البسيط

Simple linear Regression

محتويات الفصل :

الفصل الثالث

الانحدار الخطى البسيط

Simple linear Regression

(١) مقدمة :

تناولنا من قبل موضوع الارتباط سواء أكان بين ظاهرتين فقط (ارتباط بسيط) أو بين عدة ظواهر (ارتباط متعدد) وذلك بغرض دراسة شكل و طبيعة ودرجة العلاقة بين هذه الظواهر . لكن فى كثير من الحالات نكون أكثر اهتماماً بشكل العلاقة التى تربط بين الظاهرتين من حيث كونها علاقة خطية (على شكل خط مستقيم) أو علاقة غير خطية (على شكل منحنى) وذلك بغرض التنبؤ بقيمة إحدى الظاهرتين ، إذا علمت قيمة الظاهرة الأخرى والبحث فى شكل العلاقة بين الظواهر هو محور موضوع الانحدار .

والارتباط يختلف عن الانحدار ، فالارتباط هو دراسة التوزيع المشترك لمتغيرين أو أكثر كل يأخذ قيمة المختلفة دون تدخل من الباحث أى أن ازواج القيم المتناظرة هى قيم عشوائية ، أما الانحدار فيفترض وجود متغير واحد تابع dependent وان القراءات المسجلة عنه عشوائية ويتأثر هذا المتغير بواحد او أكثر من المتغيرات المستقلة (independent) حيث يتدخل الباحث بتثبيت قيم المتغير (أو المتغيرات) المستقلة عند مستويات معينة – او كما يقال أحياناً أنها تقاس دون خطأ – ويشاهد ويسجل القيم التى يأخذها المتغير التابع عند المستويات المحددة للمتغير (أو المتغيرات) المستقلة ، ومن ازواج القيم المتناظرة فى الحالة الأولى ، يمكن تقدير معامل الارتباط الذي يستدل منه على اتجاه ودرجة العلاقة بين المتغيرات موضوع الدراسة ، وفى الحالة الثانية يمكن تعيين الدالة التى تحدد العلاقة بين المتغير التابع والمتغيرات المستقلة وتستخدم هذه

الدالـــة فـــى التنـــبؤ بالقيم المستقبلية للمتغير التابع بدلالة قيم جديدة للمتغيرات المستقلة .

يتضـــح ممـــا ســـبق أنـــه عند قياس علاقة الارتباط بين الظواهر او المتغيرات ، لا نهتم بالبحث فى أى المتغيرات تابع وأيها مستقل أو أيهما السبب وأيهمـــا النتـــيجة ، بيـــنما هـــذه الـــتفرقة لها أهمية خاصة عند دراسة العلاقة الانحداريـــة بين المتغيرات – كما سنرى فيما بعد – وعلى ذلك يمكن القول بأن الارتباط هو وسيلة إحصائية تلخص العلاقة بين الظاهرتين (أو أكثر) فى صورة رقـــم (قيمته تتراوح بين –١، +١) ، أما الانحدار فهو وسيلة إحصائية تلخص العلاقة بين الظاهرتين أو أكثر فى صورة معادلة رياضية (قد تكون معادلة خط مســـتقيم او معادلـــة منحـــنى) بغرض تقدير قيمة المتغير التابع إذا علمت قيمة المتغير (او المتغيرات) المستقل .

ويعتبر سير فرانسيس جالتون Sir Frances Galton أول من تعرض لموضـــوع الانحـــدار فى نهاية القرن التاسع عشر (١٨٨٥) ، فقد استخدم كلمة Regression لأول مـــرة عـــندما كان يدرس العلاقة بين أطوال مجموعة من الآباء والأبناء عددهم حوالى ١٠٠٠ وقد كشفت دراسته عن نتائج هامة ملخصها ما يلي :

١- أن الآباء طوال القامة لهم أبناء طوال القامة ، والآباء قصار القامة لهم أبناء قصار القامة.

٢- متوسط الطول للأبناء من مجموعة آباء طوال القامة أقل من أطوال آبائهم ، ومتوســـط الطـــول للأبناء من مجموعة آباء قصار القامة أكبر من أطوال آبائهم .

وقـــد استخدم جالتون اصطلاح Regression (وترجمتها فى القواميس هـــى الارتـــداد أو الرجوع للخلف) قاصداً به الملاحظات السابقة . وقد ظل هذا

الاصـــطلاح مستخدماً حتى وقت قريب بواسطة أغلبية من الإحصائيين ، ولكن قاصـــدين بها الخط المرسوم لمجموعة من النقط يمثل الاتجاه العام لها بخلاف المعـــنى الذي كان يقصده جالتون من استخدام هذا الاصطلاح . ولذلك نجد ان بعـــض الإحصائيين فى السنوات الأخيرة تفضل استخدام اصطلاح خط التقدير بدلاً من خط الانحدار .Estimating line instead of Regression line

(٣) : طرق الحصول على خط الانحدار :

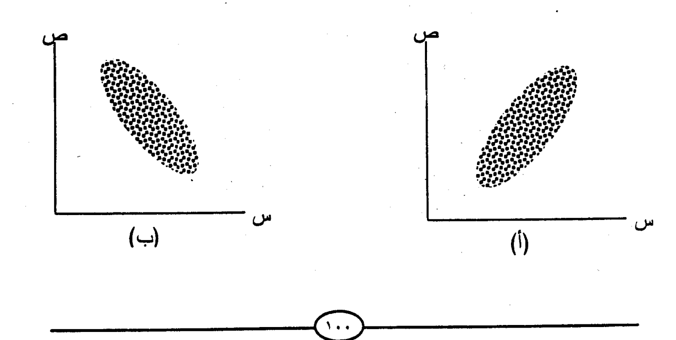
هناك طريقتين أساسيتين يمكن استخدام إحداهما للوصول إلى خط الانحدار :

١- طريقة الشكل الانتشارى (الطريقة البيانية) .

٢- طريقة المربعات الصغرى .

(٣-١) طريقة الشكل الانتشارى : Scatter Diagram Method

من الممكن اكتشاف شكل العلاقة الفعلية التى تربط بين المتغيرين (س ، ص) بواســطة الشــكل الانتشــارى ، وذلك بتمثيل المتغير المستقل (س) على المحور الأفقى والمتغير التابع (ص) على المحور الرأسي . وعند رصد أزواج المشــاهدات الفعلــية لكل من المتغيرين (س ، ص) على لوحة بيانية يمكن أن نحصــل علــى شــكل واحد لهذه البيانات من بين عدة أشكال مختلفة موضحه بالأشكال رقم أ ، ب ، ج ، د ، هـ فى شكل (١) .



ص ص

س (ج) س (د)

ص

الشكل (١)

س

(هـ)

ولإيضاح مدلول المتغير المستقل والمتغير التابع نقدم بعض الأمثلة التالية :

العلاقة بين كمية الإنتاج من القطن وكمية السماد المستخدم ، هي علاقة بين متغير تابع ومتغير مستقل ، فكمية الإنتاج تتحدد تبعا لكمية السماد المستخدم ، وبمعنى آخر بتغيير كمية السماد زيادة أو نقص يتبع ذلك تغير في كمية الإنتاج زيادة أو نقص ، أي أن التغير في كمية الإنتاج يعتمد ويتوقف على التغير الذي يحدث في كمية السماد ، لذا يقال أن السماد متغير مستقل لأن الباحث هو الذي يتحكم في اختيار مستويات أو كميات السماد المستخدمة ، أما كمية الإنتاج فهو متغير تابع لا يمكن أن يتدخل الباحث في تحديد كميته . كذلك العلاقة بين الدخل وحجم الإنفاق هي علاقة بين متغير مستقل ومتغير تابع ، فحجم الإنفاق يعتمد ويتوقف على حجم الدخل ، أيضا العلاقة بين معدلات وفيات

الأطفــال الرضع ومستوى معيشة الأسرة ، هي علاقة بين متغير تابع (معدلات الوفاة بين الأطفال الرضع) ومتغير مستقل (مستوى المعيشة).

وكمــا بينا من قبل نجد أنه إذا وقعت جميع النقط على استقامة واحدة ، أي كونت خط مستقيم ، فهذا دليل على وجود ارتباط تام بين المتغيرين (سواء أكــان طردي يصعد إلى أعلى أو سلبي يهبط إلى أسفل) أو قد تقع جميع النقط علـــى خــط منحـــنى ، مما يعني وجود ارتباط تام بينهما . لكن التمثيل الفعلي للبــيانات المستمدة من متغيرات أو ظواهر اقتصادية وتجارية لا تحقق مثل هذا الارتباط التام ، فالارتباط التام في مثل هذه الظواهر هو حالة نادرة ، وعلى ذلك مــن المــتوقع أن نجد أن التمثيل الفعلي لازواج القراءات (س ، ص) ينتشر ويتوزع حول خط مستقيم أو حول منحنى ، مما يعني وجود حالة من الارتباط غــير الــتام ، وكلما اقتربت النقط من الخط المستقيم أو من المنحنى كلما ازداد الارتــباط بيــن الظاهرتين . والخط الذي تنتشر حوله هذه النقط سواء كان خط مســتقيم أو منحنى يسمى بخط الانحدار Regression Line وهذا الخط يمكن التعبير عنه بمعادلة رياضية سواء بمعادلة خط مستقيم على الصورة :

ص = أ + ب س أو بمعادلة منحنى على الصورة ص = أ + ب س + ج س ٢

ولما كان هذا الخط ذو أهمية وفائدة كبيرة ، لأنه يستخدم في التنبؤ بقيمة المتغير التابع إذا علمت قيمة المتغير المستقل ، كما أنه يستخدم أيضا في تحديد نســبة تأثير المتغير المستقل على المتغير التابع ، فإن هذا الخط يمكن أن نصل إلــيه باحدى طريقتين : طريقة التمهيد باليد أو طريقة المربعات الصغرى . فإذا اســتخدما طريقة التمهيد باليد وهي طريقة تعتمد على المهارة الشخصية للباحث وعلى خبرته في رسم الخطوط البيانية وعلى فهمه لطبيعة البيانات محل الدراسة ، فإنــه يجب مراعاة النقاط التالية لكي نوفق في رسم أفضل خط انحدار سواء أكان خط مستقيم أو منحنى :

١- يجب أن يكون هذا الخط قريب جدا من جميع النقط .

٢- يجب أن يقع على جانبي هذا الخط عدد متساوي من النقط بقدر المستطاع .

٣- يجـــب أن تكون النقط الواقعة على جانبي هذا الخط على مسافات متساوية مـــنه ، بمعنى أن الانحرافات الموجبة (للنقط التي فوق الخط) تتعادل مع الانحرافات السالبة (للنقط التي تحت الخط) .

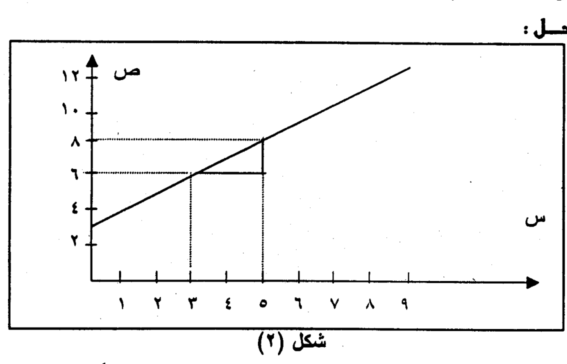
مثال (١) :

أعطيـــت البيانات التالية عن المتغيرين (س ، ص) حيث س : متغير مستقل ، ص : متغير تابع

س	٢	٣	٥	٦	٨	٩
ص	٦	٥	٧	٨	١٢	١١

استخدام طريقة الشكل الإنتشاري في استطلاع شكل العلاقة بين المتغيرين ، ثم وفق أفضل خط تراه يمثل هذه البيانات .

الحـــل :



شكل (٢)

من الواضح أن الشكل الانتشاري للنقط عند رصدها بيانياً وفق المعايير الســـابقة ، يعبر عن وجود علاقة خطية شبه مستقيمة ، يمكن التعبير عنها بخط مستقيم على الصورة :

ص = أ + ب س ، حيث (أ) عبارة عن الجزء المقطوع من المحور الرأسي وهـو يسـاوي ٣ تقريبا ، أما (ب) فهي عبارة عن ميل خط الانحدار أي ظل الزاوية التي يصنعها هذا الخط مع الاتجاه الموجب لمحور السينات أو هو مقدار التغير الذي يحدث في (ص) عندما تتغير (س) بوحدة واحدة ، أي :

$$ب = \frac{ص_٢ - ص_١}{س_٢ - س_١} = \frac{٨ - ٦}{٥ - ٣} = ١ \text{ تقريبا}$$

وعلى ذلك تصبح معادلة خط الانحدار المستقيم ص = أ + ب س على الصورة :

ص = ٣ + س

ويمكـن استخدام المعادلة السابقة في التنبؤ بقيمة المتغير التابع (ص) عـند قيمة معلومة للمتغير المستقل (س) ، فمثلا إذا كانت س = ١٠ فان قيمة (ص) المقدرة =٣+١× ١٠=١٣

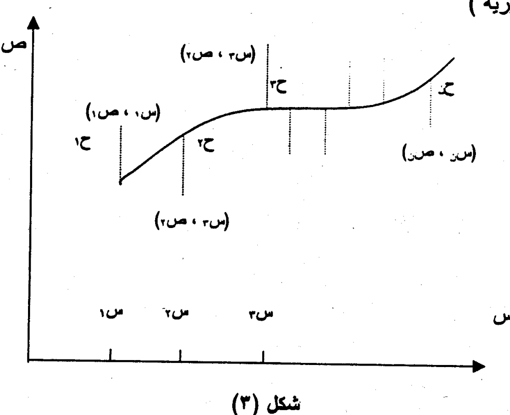
يتضـح ممـا سـبق أن التمهيد باليد طريقة سهلة وسريعة ، لكن بتعدد الأشـخاص تتعدد الخطوط البيانية لنفس البيانات ، لأن كل واحد يرى أن الخط الـذي رسـمه للبيانات هو أفضل الخطوط من وجهة نظره ، وبالتالي ينتج عن تعـدد الخطوط المستقيمة تعدد القيم المتنبأ بها للمتغير التابع (ص) عند قيمـة معينة معلومـة للمتغير المستقل (س) ، ذلك لأن كل خط ستختلف فيه قيـمة (أ ، ب) عـن الخطوط الأخرى ، وتعدد القيم التقديرية هذه دليل على عدم الدقة في توفيق معادلة تمثل هذه البيانات .

(٣-٢) طريقة المربعات الصغرى Least Square Method

لكـي نـتلافى أثر التقدير الشخصي عند رسم خط الانحدار ، نلجأ إلى طـريقة أخـرى تحكمها قواعد رياضية ثابتة لا يختلف تفسيرها أو اتباعها من شـخص لآخر . لكن دعنا في البداية نوضح المقصود باصطلاح أفضل توفيق

للخـــط البيانـــي Best Fitting Line (سواء خط مستقيم أو خط منحنى) . لنفرض أنه عند التمثيل البياني لانواج القراءات (س ، ص) حصلنا على الشكل رقم (٣) وفيه تم رسم خط بياني يتوسط النقط الفعلية بقدر الامكان .

- يلاحــظ أنـــه عند احدى قيم المتغير المستقل س ولتكن $س_١$ ، أن القيمة الفعلـــية المناظرة لها وهي $ص_١$ تبتعد عن الخط البياني المرسوم بالمسافة $ح_١$ ، كذلك نجد أنه عند القيمة $س_٢$ تبتعد القيمة الفعلية المناظرة لها وهي $ص_٢$ ، عن الخط البياني بالمسافة $ح_٢$ وهكذا لباقي النقط . أي أن هذه الفروق أو الانحرافات (وتسمى بالخطأ العشوائي كما سيتضح فيما بعد) عبارة عن الفرق بين القيمة الفعلية (ص) والقيمة المقابلة لها على الخط البياني ولتكن $\hat{ص}$ (أي القيمة التي تقـــرأ مـــن علـــى الخط البياني وتسمى بالقيمة النظرية أو المتوقعة أو بالقيمة التقديرية)



شكل (٣)

وهـــذه الانحـــرافات بعضها موجب (إذا كانت القيمة الفعلية (ص) تقع أعلـــى الخط البياني) وبعضها سالب (إذا كانت القيمة الفعلية (ص) تقع أسفل

الخــط البيانــي) وبعضها صفر (إذا كانت القيمة الفعلية (ص) تقع على الخط البياني) وأفضل خط بياني يوفق في التعبير عن هذه البيانات ويمثلها بدقة كافية هو الخط الذي يجعل مجموع مربعات الانحرافات السابقة أقل ما يمكن أي :

$ح^{٢}_{١} + ح^{٢}_{٢} + ح^{٢}_{٣} + ...+ ....+ ح^{٢}_{ن}$ أقل ما يمكن

أي مجـ $ح^{٢}_{ر}$ أقل ما يمكن ...(١)

وعلى ذلك إذا كانت هناك مجموعة من الخطوط البيانية ، يمكن رسمها للمتغيريــن (س،ص) وكانت إحدى هذه الخطوط تحقق الشرط السابق (مجموع مــربعات انحرافات القيم الفعلية عن القيم التقديرية أقل ما يمكن) فإن هذا الخط يعتبر أفضل هذه الخطوط Best Line ويسمى بخط المربعات الصغرى .

ولكي نصل إلى الخط المربعات الصغرى سواء أكان خط مستقيم أو خط منحــنى ، فان طريقة المربعات الصغرى تحدد لنا الأسلوب الرياضي الذي يتبع فــي تحديــد ثوابت كل معادلة ، وسوف نتناول هذا الأسلوب مرة عندما تكون هناك علاقة خط مستقيم بين الظاهرتين ، ومرة أخرى عندما تكون هناك علاقة منحنى من الدرجة الثانية بين الظاهرتين ، وأخيراً عندما تكون هناك علاقة خط انحدار متعدد والذي يربط بين أكثر من متغيرين .

تعــود طــريقة المربعات الصغرى إلى عالم الرياضيات الألماني كارل فــريدريك جاوس . وفي ظل فروض معينة ، تتمتع تقديرات طريقة المربعات الصـــغرى بعدة خصائص احصائية هامة ، جعلتها من أدق وأفضل الطرق التي يمكن أن تستخدم في تحليل الانحدار . ولتوضيح هذه الطريقة ، نستعرض أولا الفروض التي وضعها جاوس . بفرض أنه لدينا متغير مستقل (س) ومتغير تابع (ص) يرتــبطان بعلاقة خطية على الصورة : ص = $\gamma + \beta$ س . وحيث أنه من غير المتوقع – من الناحية العملية – أن تقع جميع النقط تماما على خط

مسـتقيم واحـد ، فإن تلك العلاقة الخطية يجب أن تعدل كي تضم حد إضافي يسمى بالحد العشوائي (ψ) ، أي أن:

ص = γ + β س + ψ

ونمـوذج الانحـدار بهـذه الصورة يعتمد على عدة فروض ، بعضها يختص بالمتغـير العشوائي (ψ) والبعض الآخر يختص بالعلاقة بين المتغير العشوائي (ψ) والمتغير المستقل (س) .

فروض نموذج خط الانحدار :

١- قيمة المتغير العشوائي (ψ) قد تكون موجبة أو سالبة أو مساوية للصفر ، لكن متوسط أو توقع هذه القيم يساوى الصفر ، أي : توقع (ψ) = صفر

٢- الحدود العشوائية (أو الأخطاء العشوائية) غير مرتبطة ببعضها البعض ، أي أن تغاير الحدود العشوائية يساوي الصفر ، وهذا الفرض (سوف نتناوله فـيما بعـد) يعنـي من الناحية الفنيـة عـدم وجـود إرتبـاط ذاتـي بيـن الحـدود العشوائية (أو البواقي) أي No autocorrelation .

٣- تباين (ψ) مقدار ثابت عند جميع قيم المتغير المستقل (س) ، أي : $6^{٢}$ (ψ)= ثابـت . وهـذا الفرض (سوف نتناوله فيما بعد) يعني من الناحية الفنية وجود تجانس للتباين Homoscedasticity .

٤- قيم الأخطاء العشوائية (ψ) مستقلة (غير مرتبطة) عن قيم المتغير المستقل (س) ،أي:

تغاير (ψ ، س) = صفر

٥- المتغـير المستقل (س) مقاس بدون أخطاء ، أما المتغير التابع فقد يحتوي أو لا يحتوي على أخطاء في القياس .

٦- التوزيع الاحتمالي للمتغير (ψ) هو التوزيع الطبيعي ، توقعه وتباينه على الصورة :

توقع (ψ) = صفر ، σ^{2} (ψ) = ثابت .ويعبر عن ذلك في الصورة المختصرة:

ψ ~ م (صفر ، σ^{2} (ψ)) ، حيث تشير م إلى التوزيع الطبيعي المعتاد أو المعتدل.

٧- من نموذج الانحدار البسيط : ص = γ + β س + ψ ، نجد أن المتغير التابع ص يصبح دالة خطية في المتغير العشوائي ψ ، أي أن التوزيع الاحتمالي للمتغير (ص) يتوقف ويعتمد على التوزيع الاحتمالي للمتغير(ψ) وعلى ذلك نجد أن (ص) تتبع توزيع طبيعي معتاد بتوقع وتباين على الصورة :

توقع (ص) = γ + β س

تباين (ص) = تباين (ψ) = σ^{2} (ψ)

ويعبر عن ذلك في الصورة المختصرة: ص ~ م (γ +β س) ، σ^{2}(ψ)

لاحظ أن المتغير المستقل (س) هو متغير مثبت عند مستويات معينة ، أي أنه متغير غير عشوائي أي متغير غير احتمالي ، بمعنى أنه عند سحب العديد من العينات ، فإن قيم (س) تكون ثابتة لا تتغير من عينة لأخرى ، كذلك γ مقدار ثابت .

ونظراً لأن γ،β ، σ^{2} (ψ) هي القيم الحقيقية (أى مؤشرات) للعلاقة بين ص ، س في مجتمع الدراسة ، وهي ثوابت عادة تكون مجهولة القيمة ، لذا فإننا نلجأ إلى تقديرها من بيانات عينة عشوائية تسحب من مجتمع الدراسة . بفرض أنه سحبت عينة عشوائية حجمها (ن) مفردة ، وسجل لكل مفردة في تلك العينة قيم الظاهرتين (س ، ص) ، بذلك تصبح العلاقة بين (ص ، س) في عينة الدراسة على الصورة :

ص = أ + ب س + خ

حيـث : (خ) يناظر (Ψ) وهو يمثل الخطأ العشوائي ، أي الخطأ فـي قـيمة (ص) ، أى الفرق بين قيمة (ص) الحقيقية أي المشاهدة والمسجلة عملـياً وقيمتها المتوقعة أي الواقعة على معادلة خط الانحدار . وسوف نتناول كيفية تقدير معاملات خط الانحدار (أ ، ب) مرة عندما يكون (ص) متغير تابع أي معادلـة خط انحدار ص على س ومرة أخرى عندما يكون (س) هو المتغير التابع أي معادلة خط انحدار س على ص .

(٣) خط الانحدار المستقيم Simple Linear Regression

(٣-١) معادلة خط الانحدار ص على س (ص / س) :

إذا كانـت العلاقـة التي تربط بين المتغيرين (س ، ص) توحي بأنها تأخذ شكل خط مستقيم – كما في شكل (٢) – معادلته على الصورة :

ص = أ + ب س + خ ‎ ‎ ‎ ‎ ‎ ‎ ‎ ‎ ‎ .....(٢)

حيث ص : المتغير التابع ‎ ‎ ‎ ‎ س : المتغير المستقل

أ : مقدار ثابت ، ويمثل الجزء المقطوع من المحور الرأسي وهو عبارة عن قيمة عندما تكون قيمة س = صفر .

ب : مـيل أو معـامل خـط الانحدار أو مقدار التغير الذي يحدث في المتغير التابع (ص) عندما تتغير قيمة المتغير المستقل (س) بوحدة واحدة .

خ : الخطـأ في قيمة (ص) أي اختلاف قيمة (ص) الحقيقية عن قيمتها التقديرية والواقعة على خط الانحدار .

نعلـم أن الخـط المستقيم يتحدد تماما بعد معرفة قيم الثوابت (أ ، ب) وأن هذه الثوابت أو المجاهيل يمكن تحديدها بعدة طرق مختلفة ، لكن المهم هو اختيار الطريقة التي تحقق أفضل خط مستقيم يمثل ازواج القراءات (س ، ص)

. وفي الواقع فان طريقة المربعات الصغرى تحدد أفضل قيم للثوابت (أ ، ب) ، فمن شأن هذه القيم والتي تحدد بطريقة المربعات الصغرى أن تجعل مجموع مـــربعات انحرافات القيم الفعلية (ص) عن القيم المتوقعة أو التقديرية (صٌ^) أقـــل ما يمكن ، أي : مجـــ (ص – صٌ^)٢ أقل ما يمكن ، حيث صٌ^ هي القيم النظرية أو التقديرية أو المتوقعة أي التي تقع على خط الانحدار ، أي أنها تحققه ، أي أن صٌ^ = أ+ب س

∴ المطلوب الآن هو تحديد قيم الثوابت أ ، ب بشرط أن :

مجـــ ح٢ = مجـــ (ص – صٌ^)٢   أقل ما يمكن ، أي أن :

مجـــ ح٢ = مجـــ (ص – أ – ب س)٢ نهاية صغرى .

وبإجـــراء عمليات التفاضل الجزئي مرة بالنسبة إلى (أ) ومرة أخرى بالنسبة إلى (ب) ومساواة الناتج بالصفر نجد أن :

(6 مجـــ ح٢) / (6 أ) = ٢ مجـــ (ص – أ – ب س) (– ١) = صفر

∴ – ٢ (مجـــ ص – ن أ – ب مجـــ س) = صفر

مجـــ ص – ن أ – ب مجـــ س = صفر ، (–٢ ≠ صفر)

∴ مجـــ ص = ن أ + ب مجـــ س   ... (٣)

(6 مجـــ ح٢) / (6 ب) = ٢ مجـــ (ص – أ – ب س) (– س) = صفر

∴ – ٢ مجـــ س (ص – أ – ب س) = صفر

مجـــ س ص – أ مجـــ س – ب مجـــ س٢ = صفر ،(–٢ ≠ صفر)

∴ مجـــ س ص = أ مجـــ س + ب مجـــ س ٢   ... (٤)

وتســـمى المعـــادلات (٣) ، (٤) بالمعـــادلات المعـــتادة أو الطبيعية Normal Equations ويحل المعادلتين (٣،٤) معا بطريقة الحذف المتتالي أو

التعويض أو بالمحددات ، أو بالمصفوفات نصل إلى قيم الثوابت (أ ، ب) حيث أن : مجـ ص ، مجـ س ، مجـ س ص، مجـ س٢ هي قيم معلومة .

مـن ناحية أخرى يمكن أن نصل من المعادلتين (٣ ، ٤) إلى علاقات رياضـية للثوابـت (أ ، ب) اذا اسـتخدمنا طريقة المحددات مثلا على النحو التالي:

$$\text{مجـ ص} = \text{ن أ} + \text{ب مجـ س}$$

$$\text{مجـ س ص} = \text{أ مجـ س} + \text{ب مجـ س}^{2}$$

$$\therefore \text{أ} = \frac{\begin{vmatrix} \text{مجـ ص} & \text{مجـ س} \\ \text{مجـ س ص} & \text{مجـ س}^{2} \end{vmatrix}}{\begin{vmatrix} \text{ن} & \text{مجـ س} \\ \text{مجـ س} & \text{مجـ س}^{2} \end{vmatrix}} = \frac{\text{مجـ ص} \times \text{مجـ س}^{2} - \text{مجـ س} \times \text{مجـ س ص}}{\text{ن مجـ س}^{2} - (\text{مجـ س})^{2}} \quad ...(٥)$$

$$\text{ب} = \frac{\begin{vmatrix} \text{ن} & \text{مجـ ص} \\ \text{مجـ س} & \text{مجـ س ص} \end{vmatrix}}{\begin{vmatrix} \text{ن} & \text{مجـ س} \\ \text{مجـ س} & \text{مجـ س}^{2} \end{vmatrix}} = \frac{\text{ن مجـ س ص} - \text{مجـ س} \times \text{مجـ ص}}{\text{ن مجـ س}^{2} - (\text{مجـ س})^{2}} \quad ...(٦)$$

أو

$$\text{ب} = \frac{\text{مجـ س ص} - \text{مجـ س} \times \text{مجـ ص} / \text{ن}}{\text{مجـ س}^{2} - (\text{مجـ س})^{2} / \text{ن}} \quad ...(٧)$$

وقـد تمكـن كـل من جاوس وماركوف من اثبات أن تقديرات طريقة المـربعات الصغرى(أ، ب) هي تقديرات خطية ، وغير متحيزة ، ولها أصغر تبايـن . وهذه الخصائص الثلاث هي من أهم الخصائص الاحصائية التي يتمتع بها أي تقدير احصائي .

طريقة أخرى للوصول إلى المعادلات الطبيعية :

رأينا أن المعادلات الطبيعية (٣) ، (٤) هي الأساس في الوصول إلى الثوابت (أ ، ب) وقـد حصلنا على هذه المعادلات ، عن طريق التفاضل الجزئي للكمية مجـــ $ح^{٢}$.لكن يمكن الوصول إلى هذه المعادلات بطريقة أخرى مبسطة يسهل تذكرها إذا تتبعنا الخطوات التالية :

ص = أ + ب س (المعادلة الأصلية)

الخطوة الأولى :

اجمــع طرفي المعادلة الأصلية السابقة (ن) من المرات ، وهي حجم العيـنة ، أي عـدد ازواج المشـاهدات المسجلة عن الظاهرتين (س ، ص) فنحصل على أول معادلة طبيعية وهي:

مجـ ص = ن أ + ب مجـ س .... (وهي نفس المعادلة رقم ٣)

الخطوة الثانية :

اضــــرب طرفي المعادلة الأصلية في المتغير المستقل (س) لتصبح على الصورة : س ص = أ س + ب $س^{٢}$ ثم اجمع طرفي المعادلة بعد ذلك (ن) من المرات لنحصل على المعادلة التالية :

مجـ س ص = أ مجـ س + ب مجـ س ٢ ....(وهي نفس المعادلة رقم ٤)

صور أخرى بديلة للتقديرات (أ ، ب) :

مـن المعادلة (٣) نلاحظ ما يلي: مجـ ص= ن أ+ب مجـ س ،بقسمة طرفي المعادلة على (ن)

$$\frac{\text{مجـ ص}}{\text{ن}} = \text{أ} + \text{ب} \frac{\text{مجـ س}}{\text{ن}} \quad \text{أى} \quad \bar{\text{ص}} = \text{أ} + \text{ب}\ \bar{\text{س}}$$

$$\therefore \text{أ} = \bar{\text{ص}} - \text{ب}\ \bar{\text{س}} \quad \text{... (٨)}$$

ومـن المعادلة (٤) وباستخدام طريقة انحرافات القيم الفعلية لكل من س ، ص

عن أوساطهما الحسابية سَ ، صَ نجد أن :

مجـــ س ص = أ مجـــ س + ب مجـــ س٢

مجـــ (س – سَ) (ص – صَ) = أ مجـــ (س – سَ) + ب مجـــ (س – سَ) ٢

(س – سَ)٢ = صفر + ب مجـــ (س – سَ) ٢

[مجـــ(س –سَ)=صفر لأن مجموع انحرافات القيم عن وسطها الحسابي= صفر]

ب = مجـــ (س – سَ) (ص – صَ) / مجـــ (س – سَ)٢ ... (٩)

مثال (٣)

مستخدما بيانات مثال (١) ، أوجد معادلة خط انحدار ص على س بطريقة المربعات الصغرى.

الحل :

معادلة خط انحدار ص على س (ص / س) هي :

ص = أ + ب س

ويمكـــن تحديد الثوابت (أ ، ب) بطريقة المربعات الصغرى إما بحل مجموعة المعادلات الطبيعية معا أو باستخدام العلاقات الرياضية وكلاهما يعطي نفس النتيجة ، وسوف نعرض للحلين بغرض التوضيح :

جدول (١)

س	ص	س ص	س٢	س٢
٢	٦	١٢	٤	٣٦
٣	٥	١٥	٩	٢٥
٥	٧	٣٥	٢٥	٤٩
٦	٨	٤٨	٣٦	٦٤
٨	١٢	٩٦	٦٤	١٤٤
٩	١١	١٩	٨١	١٢١
٣٣	٤٩	٣٠٥	٢١٩	٤٣٩

تحديد (أ ، ب) باستخدام المعادلات الطبيعية :

ص = أ + ب س

مجـ ص = ن أ + ب مجـ س

مجـ س ص = أ مجـ س + ب مجـ $س^{٢}$

بالتعويض بالمجاميع الموضحة بجدول (١)

٤٩ = ٦ أ + ٣٣ ب

٣٠٥ = ٣٣ أ + ٢١٩ ب

بضرب المعادلة الأولى في (٣٣) وبضرب المعادلة الثانية في (٦) ثم الطرح نجد أن :

١٦١٧ = ١٩٨ أ + ١٠٨٩ ب

١٨٣٠ = ١٩٨ أ + ١٣١٤ ب

-٢١٣ = صفر – ٢٢٥ ب

ب = $\frac{٢١٣}{٢٢٥}$ = ٠,٩٥

وبالتعويض عن قيمة (ب) في أي معادلة ولتكن الأولى نصل إلى قيمة (أ)

٤٩ = ٦ أ + ٣٣ × ٠,٩٥ = ٦ أ + ٣١,٣٥

٤٩ – ٣١,٣٥ = ٦ أ ١٧,٦٥ = ٦ أ أ = $\frac{١٧,٦٥}{٦}$ = ٢,٩٤

معادلة خط الانحدار ص = أ + ب س بعد تحديد الثوابت أ ، ب تصبح على الصورة :

ص = ٢,٩٤ + ٠,٩٥ س

ويلاحـــظ أن المعادلة التي توصلنا إليها بالطريقة البيانية من قبل قريبة جدا من هذه المعادلة .

تحديد (أ ، ب) باستخدام العلاقات الرياضية :

ص = أ + ب س ، حيث :

$$ب = \frac{\text{مجـ س ص} - \frac{\text{مجـ س} \times \text{مجـ ص}}{ن}}{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{ن}} = \frac{٣٠٥ - \frac{٣٣ \times ٤٩}{٦}}{٢١٩ - \frac{(٣٣)^{٢}}{٦}}$$

$$ب = \frac{٣٠٥ - ٦٩,٥}{٢١٩ - ١٨١,٥} = \frac{٣٥,٥}{٣٧,٥} = ٠,٩٥$$

$$أ = \bar{ص} - ب\,\bar{س} = \frac{\text{مجـ ص}}{ن} - ٠,٩٥ \times \frac{\text{مجـ س}}{ن}$$

$$= \frac{٤٩}{٦} - ٠,٩٥ \times \frac{٣٣}{٦} = ٨,١٧ - ٠,٩٥ \times ٥,٥$$

$$= ٨,١٧ - ٥,٢٣ = ٢,٩٤$$

∴ معادلة خط الانحدار هى :

ص = ٢,٩٤ + ٠,٩٥ س وهى نفس المعادلة السابقة .

<u>صور أخرى لمعاملات خط الانحدار بطريقة الانحرافات البسيطة والمعدلة :</u>

كمـا ذكـرنا من قبل ، يمكن تبسيط العمليات الحسابية اللازمة لتقدير الثوابت (أ، ب) وخاصة (ب) عن طريق اختيار أوساط فرضية من كل من س ، ص ، ولتكن $و_{١}$ ، $و_{٢}$.

$$∴ ب = \frac{\text{مجـ س ص} - \frac{\text{مجـ س} \times \text{مجـ ص}}{ن}}{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{ن}}$$ ... الطريقة المباشرة

فإذا استخدمنا طريقة الانحرافات البسيطة نجد أن :

$$ب = \frac{\text{مجـ } ح_{١} ح_{٢} - \frac{\text{مجـ } ح_{١} \times \text{مجـ } ح_{٢}}{ن}}{\text{مجـ } ح_{١}^{٢} - \frac{(\text{مجـ } ح_{١})^{٢}}{ن}}$$

حيــث $\text{ح}_1 = \text{س} - \text{و}_1$ ، $\text{ح}_2 = \text{ص} - \text{و}_2$ ، أمــا الأرقــام (١ ، ٢) فــترمز للمتغــيرات (س ، ص) على التوالى . وإذا كانت الانحرافات البسيطة ح_1 ، ح_2 كل يقبل القسمة على عامل مشترك ث_1، ث_2 على التوالى فإن :

$$\text{ب} = \frac{\text{ث}_1\,\text{ث}_2\left[\text{مجـ}\,\bar{\text{ح}}_1\,\bar{\text{ح}}_2 - (\text{مجـ}\,\bar{\text{ح}}_1 \times \text{مجـ}\,\bar{\text{ح}}_2)\,/\,\text{ن}\right]}{\text{ث}_1^2\left[\text{مجـ}\,\bar{\text{ح}}_1^2 - (\text{مجـ}\,\bar{\text{ح}}_1)^2\,/\,\text{ن}\right]}$$

$$= \left(\frac{\text{ث}_2}{\text{ث}_1}\right)\frac{\left[\text{مجـ}\,\bar{\text{ح}}_1\,\bar{\text{ح}}_2 - (\text{مجـ}\,\bar{\text{ح}}_1 \times \text{مجـ}\,\bar{\text{ح}}_2)\,/\,\text{ن}\right]}{\left(\text{مجـ}\,\text{ح}_1^2 - (\text{مجـ}\,\bar{\text{ح}}_1)^2\,/\,\text{ن}\right)} \quad \ldots (11)$$

أما الثابت (أ) فيكفى استخدام المعادلة رقم (٨) حيث :

$$\text{أ} = \bar{\text{ص}} - \text{ب}\,\bar{\text{س}} = \frac{\text{مجـ ص}}{\text{ن}} - \text{ب} \times \frac{\text{مجـ س}}{\text{ن}}$$

مثال (٣):

الجــدول الــتالى يوضـــح كميات الإنتاج (ص) من محصول القمح الناتجة عن كميات مختلفة من السماد (س) فى عينة من ١٠ أفدنة :

كمية الإنتاج :ص	٤٠	٤٤	٤٦	٤٨	٥٢	٥٨	٦٠	٦٨	٧٤	٨٠
كمية السماد :س	٦	١٠	١٢	١٤	١٦	١٨	٢٢	٢٤	٢٦	٣٢

المطلــوب توفــيق أفضــل معادلــة خط انحدار ص على س بطريقة المربعات الصغرى ثم ارسم هذا الخط بيانياً – ما هى كمية الإنتاج المتوقعة إذا كانت كمية السماد = ٢٠ وحدة .

الحل :

بدلاً من استخدام القيم الحقيقية لكل من س ، ص فى إيجاد معاملات خط الانحــدار أ ، ب يمكن استخدام طريقة الانحرافات البسيطة وذلك بطرح وسط فرضى (و_1) من قيم (س) وليكن $\text{و}_1 = ٢٠$ وبطرح وسط فرضى (و_2) من قيم (ص) وليكــن $\text{و}_2 = ٥٠$ ، بذلك نحصل على الانحرافات ح_1 ، ح_2 حيث ح_1 =

س – ٢٠ ، $ح_٢$ = ص – ٥٠ وبالتالى بدلاً من إجراء العمليات الحسابية السابقة على (س ، ص) تجرى على بدائلها $ح_١$ ، $ح_٢$.

جدول (٢)

س	ص	$ح_١$	$ح_٢$	$ح_١^٢$	$ح_٢^٢$	$ح_١ ح_٢$
٦	٤٠	-١٤	-١٠	١٩٦	١٠٠	١٤٠
١٠	٤٤	-١٠	-٦	١٠٠	٣٦	٦٠
١٢	٤٦	-٨	-٤	٦٤	١٦	٣٢
١٤	٤٨	-٦	-٢	٣٦	٤	١٢
١٦	٥٢	-٤	+٢	١٦	٤	-٨
١٨	٥٨	-٢	+٨	٤	٦٤	-١٦
٢٢	٦٠	+٢	+١٠	٤	١٠٠	٢٠
٢٤	٦٨	+٤	+١٨	١٦	٣٢٤	٧٢
٢٦	٧٤	+٦	+٢٤	٣٦	٥٧٦	١٤٤
٣٢	٨٠	+١٢	+٣٠	١٤٤	٩٠٠	٣٦٠
١٨٠	٥٧٠	-٢٠	+٧٠	٦١٦	٢١٢٤	٨١٦

معادلة خط الانحدار ص / س : ص = أ + ب س حيث :

$$ب = \frac{\text{مجـ } ح_١ ح_٢ - \frac{\text{مجـ } ح_١ \times \text{مجـ } ح_٢}{ن}}{\text{مجـ } ح_١^٢ - \frac{(\text{مجـ } ح_١)^٢}{ن}}$$

$$= \frac{٨١٦ - \frac{-٢٠ \times ٧٠}{١٠}}{٦١٦ - \frac{(-٢٠)^٢}{١٠}} = \frac{٨١٦ + ١٤٠}{٦١٦ - ٤٠} = \frac{٩٥٦}{٥٧٦} = ١,٦٦$$

$$أ = \bar{ص} - ب \bar{س} = \frac{\text{مجـ ص}}{ن} - ١,٦٦ \times \frac{\text{مجـ س}}{ن}$$

$= \frac{٥٧٠}{١٠} - ١,٦٦ \times \frac{١٨٠}{١٠} = ٥٧ - ١,٦٦ \times ١٨ = ٥٧ - ٢٩,٨٨ = ٢٧,١٢$

∴ معادلة خط الانحدار هى :

ص = ٢٧,١٢ + ١,٦٦ س

تقدير كمية الإنتاج عندما تكون كمية السماد = ٢٠ وحدة هى :

$\hat{ص}$ = ٢٧,١٢ + ١,٦٦ × ٢٠ = ٢٧,١٢ + ٣٣,٢٠ = ٦٠,٣٢ وحدة

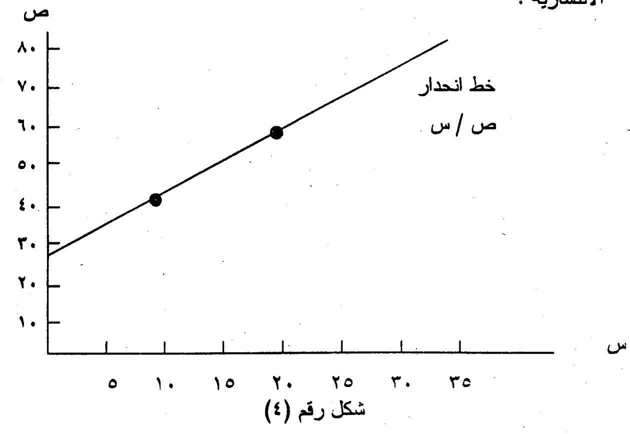
ولرســم معادلــة خــط الانحدار السابقة ، نحتاج إلى معرفة الاحداثيات السينية والصادية لنقطتين فقط . فمثلاً :

عندما تكون س = ١٠ فإن ص = ٢٧,١٢ + ١,٦٦ × ١٠ = ٤٣,٧٢ .

وعندما تكون س = ٢٠ فإن ص = ٢٧,١٢ + ١,٦٦٦ × ٢٠ = ٦٠,٣٢

∴ النقطتان هما : (١٠ ، ٤٣,٧٢) ، (٢٠ ، ٦٠,٣٢)

فــي الشكل رقم (٤) تم رصد ازواج المشاهدات (النقط الانتشارية) وتم رســم خــط الانحدار التقديري ومن الواضح أن هذا الخط يقترب جداً من النقط الانتشارية .



شكل رقم (٤)

ملاحظات :

١- يلاحظ أن الأوساط الحسابية لكل من س ، ص أعداد صحيحة ($\bar{س}$ = ١٨ ، $\bar{ص}$ = ٥٧) ومن ثم يكون من الأفضل اختيار الأوسط الفرضية $و_١$ ، $و_٢$ مساوية للأوساط الحسابية $\bar{س}$، $\bar{ص}$ ، وفى هذا تخفيض كبير فى حجم العمليات الحسابية ، حيث نستخدم المعادلة رقم (٩) لحساب قيمة (ب) ، أما (أ) فلا خلاف فى القانون الرياضي المستخدم لها وهو دائماً العلاقة رقم(٨).

٢- يلاحظ أيضاً أن الانحرافات البسيطة $ح_١$ ، $ح_٢$ فى هذا المثال كل منها يمكن تبسيطه أكثر من ذلك بالقسمة على عامل مشترك $ث_١$ ، $ث_٢$ حيث نجد أن $ث_١$ = ٢ ، $ث_٢$ = ٢ (ليس من الضرورى أن تكون $ث_١$ = $ث_٢$) وفى هذا أيضا تخفيض كبير فى حجم العمليات الحسابية ، حيث نستخدم المعادلة رقم (١١) لحساب قيمة (ب) .ويترك للقارئ إعادة الحل مرة أخرى مراعياً الملاحظات (١ ، ٢) ليتأكد أنه سيصل حتما لنفس معادلة خط الانحدار السابقة .

(٣-٣) معادلة خط انحدار س على ص (س/ص) :

هناك بعض المتغيرات تتبدل فيها طبيعة المتغير المستقل والمتغير التابع، بمعني قد يصبح المتغير المستقل هو المتغير التابع أو العكس . فمثلاً العلاقة بين الطول والوزن قد يكون أحدهما هو التابع والآخر هو المستقل .

إذا افترضنا أن هناك علاقة خطية من الدرجة الأولى تربط ما بين المتغير المستقل (ص) ، وهو المتغير الذى تتحدد قيمة مسبقاً دون تدخل من الباحث ، والمتغير التابع (س) وهو المتغير الذي تتحدد قيمة تبعاً للتغير الذي يحدث فى المتغير المستقل (ص) فإن معادلة خط الانحدار س على ص (س/ص) تكون على الصورة :

س = جـ + د ص ... (١٢)

حيث : ص : متغير مستقل ، س : متغير تابع .

جـ : مقدار ثابت ، وهو عبارة عن الجزء المقطوع من المحور الأفقي (محور السينات) أو هو عبارة عن قيمة (س) عندما تكون ص = صفر .

د : ميل أو معامل خط انحدار س على ص وهو عبارة عن مقدار التغير الذي يحدث في المتغير التابع (س) إذا تغير المتغير المستقل (ص) بوحدة واحدة.

وأفضل تقدير للثوابت (جـ ، د) هي تلك التي نحصل عليها بطريقة المربعات الصغرى ، وهى الطريقة التي تجعل مجموع مربعات انحرافات القيم الفعلية للمتغير التابع ($\hat{س}$) عن القيم المتوقعة (س) والواقعة على خط الانحدار أقل ما يمكن . ويلاحظ ان الانحرافات هنا هي انحرافات أفقية ، عكس الحال في معادلة خط انحدار (ص على س) حيث كانت الانحرافات هناك رأسية .

وبتتبع طريقة المربعات الصغرى ، سنجد أننا نهدف إلى جعل الكمية التالية :

مجـ $ح^2$ = مجـ $(س - \hat{س})^2$ أقل ما يمكن

وبما أن القيم التقديرية $\hat{س}$ تقع على خط الانحدار فهي إذن تحققها ومن ثم نجد أنه مطلوب جعل مجـ $(س - جـ - د ص)^2$ أقل ما يمكن

بإجراء عمليات التفاضل الجزئي للكمية السابقة مرة بالنسبة إلى (جـ) ومرة أخرى بالنسبة إلى (د) ومساواة الناتج في كل مرة بالصفر ، نصل إلى المعادلات الطبيعية التالية :

مجـ س = ن جـ + د مجـ ص ... (١٣)

مجـ س ص = جـ مجـ ص + د مجـ $ص^2$ ... (١٤)

وبحل المعادلتين (١٣ ، ١٤) معاً بأي طريقة جبرية نصل إلى قيمة الثوابت (جـ ، د) أو يمكن استخدام العلاقات الرياضية التالية :

∴ س = جـ + د ص .

$$\therefore \text{د} = \frac{\text{مجـ س ص} - \frac{\text{مجـ س} \times \text{مجـ ص}}{\text{ن}}}{\text{مجـ ص}^{2} - \frac{(\text{مجـ ص})^{2}}{\text{ن}}} \quad ...(١٥)$$

أو :

$$\text{د} = \frac{\text{مجـ (س} - \bar{\text{س}}\text{) (ص} - \bar{\text{ص}}\text{)}}{\text{مجـ (ص} - \bar{\text{ص}}\text{)}^{2}} \quad ...(١٦)$$

$$\bar{\text{جـ}} = \bar{\text{س}} - \text{د } \bar{\text{ص}} \quad ...(١٧)$$

$$\text{حيث : } \bar{\text{س}} = \frac{\text{مجـ س}}{\text{ن}} \text{ ، } \bar{\text{ص}} = \frac{\text{مجـ ص}}{\text{ن}}$$

وإذا استخدمنا طـريقة الانحرافات البسيطة او المعدلة بغرض تبسيط العمليات الحسابية نجد أن ميل خط الانحدار (د) يأخذ الصور التالية :

$$\text{د} = \frac{\text{مجـ ح}_{1}\text{ ح}_{2} - \frac{\text{مجـ ح}_{1} \times \text{مجـ ح}_{2}}{\text{ن}}}{\text{مجـ ح}^{2}_{2} - \frac{(\text{مجـ ح}_{2})^{2}}{\text{ن}}} \quad ...(١٨)$$

أو :

$$\text{د} = \frac{\text{ث}_{1}}{\text{ث}_{2}} \times \frac{\text{مجـ ح}'_{1}\text{ ح}'_{2} - \frac{\text{مجـ ح}'_{1} \times \text{مجـ ح}'_{2}}{\text{ن}}}{\text{مجـ ح}'^{2}_{2} - \frac{(\text{مجـ ح}'_{2})^{2}}{\text{ن}}} \quad ...(١٩)$$

حيث (ن) في جميع العلاقات السابقة هي حجم العينة ، أي عدد أزواج القراءات المتناظرة بين (س ، ص) .

ملاحظات :

١- يلاحظ أن بسط ميل خط الانحدار (ب أو د) في جميع المعادلات السابقة يساوى بسط معامل الارتباط الخطى البسيط لبيرسون .

٢- يلاحظ ان هناك تماثل وتطابق كامل بين بسطي ميل خط الانحدار (ب ، د).

(٣-٣) العلاقة بين معادلتي خط الانحدار :

بينا من قبل أن معادلتي خط الانحدار (ص على س) ، (س على ص) هما على الصورة التالية:

ص = أ + ب س ... (ص / س)

س = جـ + د ص ... (س / ص)

فإذا أردنا تقدير قيمة ص (متغير تابع) عند قيمة محددة للمتغير المستقل س ، فإننا نستخدم معادلة خط انحدار (ص/س) وإذا أردنا تقدير قيمة س (كمتغير تابع) عند قيمة محددة للمتغير المستقل ص ، فإننا نستخدم معادلة خط انحدار (س/ص) .

والمعادلتين ص = أ + ب س ، س = جـ + د ص ، يختلفان عن بعضهما اختلافاً كاملاً، ففي المعادلة الأولى يتم تقدير الثوابت (أ ، ب) بطريقة المربعات الصغرى ، بحيث تجعل الانحرافات الرأسية بين القيم الفعلية للمتغير التابع (ص) والقيم التقديرية ($\hat{ص}$) الواقعة على خط الانحدار أقل ما يمكن ، بينما في المعادلة الثانية يتم تقدير الثوابت (جـ ، د) بطريقة المربعات الصغرى أيضاً ، بحيث تجعل الانحرافات الأفقية للمتغير التابع (س) والقيم التقديرية ($\hat{س}$) الواقعة على خط الانحدار أقل ما يمكن .

وهناك علاقة رياضية تربط ما بين معادلتي خط الانحدار يمكن إيضاحها على النحو التالي :

$$\therefore ب = \frac{مجـ (س - \bar{س})(ص - \bar{ص})}{مجـ (س - \bar{س})^{2}} = \frac{مجـ س ص - مجـ س \times مجـ ص / ن}{مجـ س^{2} - (مجـ س)^{2} / ن}$$

$$\therefore \text{د} = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})}{\text{مجـ}(\text{ص} - \bar{\text{ص}})^{٢}} = \frac{\text{مجـ س ص} - \text{مجـ س} \times \text{مجـ ص} / \text{ن}}{\text{مجـ ص}^{٢} - (\text{مجـ ص})^{٢} / \text{ن}}$$

$$\therefore \text{ب} \times \text{د} = \frac{[\text{مجـ}(\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})]^{٢}}{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢} \times \text{مجـ}(\text{ص} - \bar{\text{ص}})^{٢}}$$

$$= \frac{\left(\text{مجـ س ص} - \frac{\text{مجـ س} \times \text{مجـ ص}}{\text{ن}}\right)^{٢}}{\left(\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{\text{ن}}\right)\left(\text{مجـ ص}^{٢} - \frac{(\text{مجـ ص})^{٢}}{\text{ن}}\right)}$$

ونـــاتج حاصل الضرب السابق ما هو إلا مربع معامل الارتباط الخطي البســـيط لبيرسون أما الجذر التربيعي لحاصل ضرب (ب × د) فيعطى معامل الارتباط الخطي (ر) . نخلص من هذا إلى أن :

$$\boxed{\text{ر} = \sqrt{\text{ب} \times \text{د}}} \qquad ... (٢٠)$$

أي أن الجــذر التربيعي لحاصل ضرب معاملي خط الانحدار (ب ، د) يعطـــى معـــامل الارتباط (ر) بين الظاهرتين (س ، ص) وهذه علاقة إضافية لحساب معامل الارتباط العادي .

ملاحظات :

١- يلاحظ عند رسم خط الانحدار (ص/س) ، (س/ص) في شكل واحد نجد ، أنهمـــا يتقاطعان عند النقطة ($\bar{\text{س}}$ ، $\bar{\text{ص}}$) ، بمعنى أن الإحداثي الأفقي لنقطة التقاطع هو الوسط الحسابي للظاهرة (س) والإحداثي الرأسي لنقطة التقاطع هو الوسط الحسابي للظاهرة (ص) ويرجع ذلك إلى النقطة ($\bar{\text{س}}$،$\bar{\text{ص}}$) تحقق كلتا المعادلتين اى تقع على خطى الانحدار .

٢- يلاحظ أنه إذا كان هناك ارتباط طردي موجب بين الظاهرتين (س ، ص) ، فإن خطي الانحدار (ص/س) ، (س/ص) يميل كل منهما على المحور الأفقـــي ميلاً موجباً أي أن كل منهما يصنع زاوية حادة مع محور السينات (المحور الأفقي) .

٣- إذا كان الارتباط بين الظاهرتين ارتباط عكسي سالب ، فإن خطى الانحدار يصنع كل منهما زاوية منفرجة مع المحور الأفقي .

٤- أمـــا إذا كان هناك ارتباط تام موجب أو سالب بين الظاهرتين ، فإن خطى الانحدار (ص/س) (س/ص) ينطبقان على بعضهما البعض .

٥- وإذا كان الارتباط بين الظاهرتين منعدم (ر = صفر) ، فإن خطى الانحدار يتعامدان مع بعضهما ويصنعا بينهم زاوية قائمة = ٩٠ ° .

مثال (٤)

الجدول التالى يوضح أوزان عينة مكونة من ١٢ أب (س) وأكبر الأبناء (ص) .

(س)	٦٥	٦٣	٦٧	٦٤	٦٨	٦٢	٧٠	٦٦	٦٨	٦٧	٦٩	٧١
(ص)	٦٨	٦٦	٦٨	٦٥	٦٩	٦٦	٦٨	٦٥	٧١	٦٧	٦٨	٧٠

<u>المطلوب :</u>

١- أوجد معادلة خط انحدار ص على س .

٢- أوجد معادلة خط انحدار س على ص .

٣- من معادلتي خط الانحدار ، استنتج قيمة معامل الارتباط بين طول الأب وطول الابن .

٤- أعــرض بيانــياً معادلتـــي خط الانحدار بجانب النقط الانتشارية للقراءات الفعلية.

الحل :

يمكـــن استنتاج معادلتي خط الانحدار إما باستخدام القراءات الأصلية ، وفـــى هـــذا مجهود حسابي ضخم أو باستخدام انحرافات القراءات الأصلية عن الأوســـاط الحسابية لكل من س ، ص وفى هذا مجهود حسابي ضخم أيضاً لأن الأوســـاط الحســـابية هـــنا أعداد كسرية .ولكن من الأفضل استخدام انحرافات القراءات الأصلية عن أوساط فرضية و$_١$ ، و$_٢$ لكل من س ، ص فإذا فرض أن و$_١$ = و$_٢$ = ٦٠ فإننا يمكن أن نكون الجدول التالي :

١- معادلة خط انحدار ص على س : ص = أ + ب س   حيث :

$$ب = \frac{مجـ\ ح_١ ح_٢ - \frac{مجـ\ ح_١ \times مجـ\ ح_٢}{ن}}{مجـ\ ح_١^٢ - \frac{(مجـ\ ح_١)^٢}{ن}}$$

$$= \frac{٦٤٧ - \frac{٨٠ \times ٩١}{١٢}}{٦١٨ - \frac{(٨٠)^٢}{١٢}} = \frac{٦٤٧ - ٦٠٦,٦٧}{٦١٨ - ٥٣٣,٣٣} = \frac{٤٠,٣٣}{٨٤,٦٧} = ٠,٤٧٦$$

جدول (٣)

س	ص	$ح_١$	$ح_٢$	$ح_١^٢$	$ح_٢^٢$	$ح_١ ح_٢$
٦٥	٦٨	٥	٨	٢٥	٦٤	٤٠
٦٣	٦٦	٣	٦	٩	٣٦	١٨
٦٧	٦٨	٧	٨	٤٩	٦٤	٥٦
٦٤	٦٥	٤	٥	١٦	٢٥	٢٠
٦٨	٦٩	٨	٩	٦٤	٨١	٧٢
٦٢	٦٦	٢	٦	٤	٣٦	١٢
٧٠	٦٨	١٠	٨	١٠٠	٦٤	٨٠
٦٦	٦٥	٦	٥	٣٦	٢٥	٣٠
٦٨	٧١	٨	١١	٦٤	١٢١	٨٨
٦٧	٦٧	٧	٧	٤٩	٤٩	٤٩
٦٩	٦٨	٩	٨	٨١	٦٤	٧٢
٧١	٧٠	١١	١٠	١٢١	١٠٠	١١٠
٨٠٠	٨١١	٨٠	٩١	٦١٨	٧٢٩	٦٤٧

$أ = \bar{ص} - ب\ \bar{س}$

$$= \frac{\text{٨١١}}{\text{١٢}} - \text{٠,٤٧٦} \times \frac{\text{٨٠٠}}{\text{١٢}} = \text{٦٧,٥٨} - \text{٣١,٧٥} = \text{٣٥,٨٣}$$

∴ معادلة خط الانحدار ص /س هى : **ص = ٣٥,٨٣ + ٠,٤٧٦ س**

٢- معادلة خط انحدار س على ص : س = جـ + د ص   حيث :

$$\text{د} = \frac{\text{مجـ ح}_{1}\text{ح}_{2} - \dfrac{\text{مجـ ح}_{1} \times \text{مجـ ح}_{2}}{\text{ن}}}{\text{مجـ ح}_{2}^{2} - \dfrac{(\text{مجـ ح}_{2})^{2}}{\text{ن}}}$$

$$= \frac{\text{٦٤٧} - \dfrac{\text{٨٠} \times \text{٩١}}{\text{١٢}}}{\text{٧٢٩} - \dfrac{(\text{٩١})^{2}}{\text{١٢}}} = \frac{\text{٤٠,٣٣}}{\text{٧٢٩} - \text{٦٩٠,٨}} = \frac{\text{٤٠,٣٣}}{\text{٣٨,٩٢}} = \text{١,٠٣٦}$$

$$\text{جـ} = \bar{\text{س}} - \text{د}\,\bar{\text{ص}}$$

$$= \frac{\text{٨٠٠}}{\text{١٢}} - \text{١,٠٣٦} \times \frac{\text{٨١١}}{\text{١٢}} = \text{٦٦,٦٧} - \text{٧٠,٠٢} = -\text{٣,٣٥}$$

∴ معادلة خط انحدار س / ص هي : **س = − ٣,٣٥ + ١,٠٣٦ ص**

(٣) استنتاج قيمة معامل الارتباط :

∴ ص = ٣٥,٨٣ + ٠,٤٧٦ س   (ص / س)

س = − ٣,٣٥ + ١,٠٣٦ ص   (س / ص)

أي أن :   ب = ٠,٤٧٦ ، د = ١,٠٣٦

$$\therefore \text{ر} = \sqrt{\text{ب} \times \text{د}}$$

$$= \sqrt{\text{٠,٤٧٦} \times \text{١,٠٣٦}} = \sqrt{\text{٠,٤٩٣}} = \text{٠,٧}$$

ملحوظة :

بمـــا أن قيمة معامل الارتباط يجب أن تكون أقل من الواحد الصحيح ، فلابد وأن يكون أحد معاملي خط الانحدار (ب أو د) أقل من الواحد الصحيح .

(٤) معادلتي خط الانحدار بيانيا :

لرسم أي معادلة انحدارية ، تحتاج إلى معرفة احداثيات أي نقطتين اختياريتين .

* لرسم معادلة خط انحدار ص / س :

∴ ص = ٣٥,٨٣ + ٠,٤٧٦ س

اذا كانت س = ٦٠ فان ص = ٣٥,٨٣ + ٠,٤٧٦ × ٦٠ = ٦٤,٣٩

واذا كانت س = ٧٠ فان ص = ٣٥,٨٣ + ٠,٤٧٦ × ٧٠ = ٦٩,١٥

∴ النقطتان هما : (٦٠، ٦٤,٣٩) ، (٧٠ ، ٦٩,١٥) .

وهذه المعادلة موضحة بالخط رقم (١) في شكل (٥)

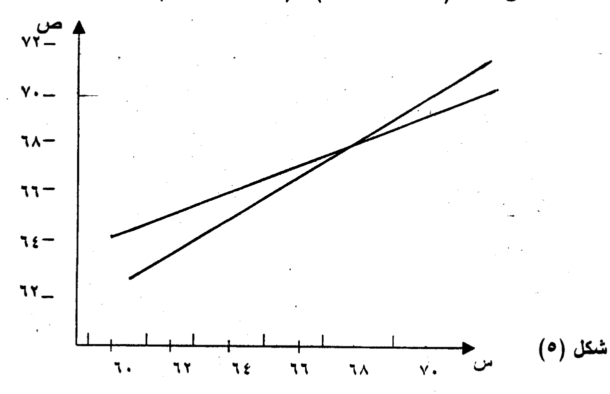
* ولرسم معادلة خط انحدار س / ص :

س = - ٣,٣٥ + ١,٠٣٦ ص .

اذا كانت ص = ٦٥ فان س = - ٣,٣٥ + ١,٠٣٦ × ٦٥ = ٦٣,٩٩

وإذا كانت ص = ٧٠ فان س = - ٣,٣٥ + ١,٠٣٦ × ٧٠ = ٦٩,١٧

∴ النقطتان هما : (٩٩، ٦٣ ، ٦٥) ، (١٧ ، ٦٩ ، ٧٠)



شكل (٥)

وهذه المعادلة موضحة بالخط رقم (٢) في شكل (٥)

ويلاحظ أن إحداثيات نقطة التقاطع هي (٦٦,٧ ، ٦٧,٦) وهذه الإحداثيات ما هي إلا الأوساط الحسابية لكل من $\bar{س}$ ، $\bar{ص}$ على التوالي أي $\bar{س}$ = ٦٦,٧ ، $\bar{ص}$ = ٦٧,٦

(٤) العلاقة بين ميل خط الانحدار ومعامل الارتباط:

١- معادلة خط انحدار ص على س (ص / س) :

$$ص = أ + ب\,س$$

$$\therefore ب = \frac{مجـ (س - \bar{س})(ص - \bar{ص})}{مجـ (س - \bar{س})^2} \quad \text{(ميل خط الانحدار)}$$

∴ معامل الارتباط الخطي البسيط لبيرسون بين الظاهرتين (س ، ص) على الصورة :

$$ر = \frac{مجـ (س - \bar{س})(ص - \bar{ص})}{\sqrt{مجـ (س - \bar{س})^2 \times مجـ (ص - \bar{ص})^2}}$$

بضرب العلاقة الأخيرة في الكمية $\frac{\sqrt{مجـ (س - \bar{س})^2}}{\sqrt{مجـ (س - \bar{س})^2}}$ نصل إلى:

$$ر = \frac{مجـ (س - \bar{س})(ص - \bar{ص})}{مجـ (س - \bar{س})^2} \times \frac{\sqrt{مجـ (س - \bar{س})^2}}{\sqrt{مجـ (ص - \bar{ص})^2}}$$

$$ر = ب \times \frac{ع_س}{ع_ص} \quad ...(٢١)$$

حيث $ع_س$ ، $ع_ص$ عبارة عن الانحراف المعياري لكل من س ، ص على التوالي . والعلاقة الأخيرة يمكن إعادة صياغتها على النحو التالي :

$$ب = ر \times \frac{ع_{ص}}{ع_{س}} \quad ...(٢٢)$$

العلاقـــات الأخـــيرة (٢١) ، (٢٢) تعنى أنه يمكن إيجاد قيمة (ر) إذا علمت قيمة (ب) والعكس صحيح . وعلى ذلك يمكن إعادة صياغة معادلة خط انحدار ص على س بدلالة معامل الارتباط (ر) على النحو التالي :

ص = أ + ب س

∴ أ = $\bar{ص}$ – ب $\bar{س}$ حيث $\bar{س}$ ، $\bar{ص}$ هما الأوساط الحسابية لكل من الظاهرتين (س ، ص) .

∴ ص = ($\bar{ص}$ – ب $\bar{س}$) + ب س

$$∴ (ص – \bar{ص}) = ب (س – \bar{س}) \quad ...(٢٣)$$

وبما أن ب = ر × $\frac{ع_{ص}}{ع_{س}}$

$$∴ (ص – \bar{ص}) = ر \times \frac{ع_{ص}}{ع_{س}} (س – \bar{س}) \quad ...(٢٤)$$

أو على الصورة التالية :

$$\frac{ص – \bar{ص}}{ع_{ص}} = ر \frac{س – \bar{س}}{ع_{س}} \quad ...(٢٥)$$

(٢) معادلة خط انحدار س على ص (س / ص) :

س = جـ + د ص .

باتباع نفس المنهج السابق يمكن أن نصل إلى النتائج التالية :

$$ر = د \times \frac{ع_{ص}}{ع_{س}} \quad ...(٢٦)$$

$$\boxed{\text{د} = \text{ر} \times \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}}} \quad ...(27)$$

كمـــا أنه يمكن إعادة صياغة معادلة خط انحدار س /ص السابقة بدلالة معلومية معامل الارتباط (ر) على النحو التالي :

$\text{س} = \text{جـ} + \text{د ص}$ ، $\text{جـ} = \bar{\text{س}} - \text{د}\ \bar{\text{ص}}$

$\therefore \text{س} = (\bar{\text{س}} - \text{د}\ \bar{\text{ص}}) + \text{د ص}$

$$\boxed{(\text{س} - \bar{\text{س}}) = \text{د}\ (\text{ص} - \bar{\text{ص}})} \quad ...(28)$$

$\therefore \text{د} = \text{ر} \times \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}}$

$$\boxed{(\text{س} - \bar{\text{س}}) = \text{ر} \times \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}}\ (\text{ص} - \bar{\text{ص}})} \quad ...(29)$$

أو

$$\frac{\text{س} - \bar{\text{س}}}{\text{ع}_{\text{س}}} = \text{ر} \times \frac{\text{ص} - \bar{\text{ص}}}{\text{ع}_{\text{ص}}} \quad ...(30)$$

أي أنه بمعلومية كل من ر ، $\bar{\text{س}}$ ، $\bar{\text{ص}}$ ، $\text{ع}_{\text{ص}}$ ، $\text{ع}_{\text{س}}$ يمكن استنتاج معادلة خط انحدار ص /س بالمعادلة (٢٤) أو معادلة خط انحدار س / ص بالمعادلة (٢٩)

مثال (٥):

أجـــرى باحـــث دراســـة تحليلية باستخدام العينة على ظاهرتين (س ، ص) وتوصل إلى المعلومات التالية عن معادلتي خط الانحدار :

٨ س – ١٠ ص + ٦٦ = صفر ، ٤ س – ١٨ ص = ٢١٤ ، تباين الظاهرة (س) = ٩

<u>المطلوب :</u>

١- ايجاد قيمة الأوساط الحسابية $\bar{س}$ ، $\bar{ص}$.

٢- معامل الارتباط بين الظاهرتين (س ، ص).

٣- الانحراف المعياري للظاهرة (ص).

الحل :

(١) إيجاد قيمة الأوساط الحسابية $\bar{س}$ ، $\bar{ص}$:

∵ معادلات الانحدار معطاة على الصورة التالية :

٨ س – ١٠ ص = – ٦٦ ، ٤٠ س – ١٨ ص = ٢١٤

وحيث أن خطي الانحدار يمرا دائما بالنقطة ($\bar{س}$ ، $\bar{ص}$) ، بمعنى أن النقطة ($\bar{س}$ ، $\bar{ص}$) تحقق كل معادلة .

∴ ٨ $\bar{س}$ – ١٠ $\bar{ص}$ = –٦٦ ، ٤٠ $\bar{س}$ – ١٨ $\bar{ص}$ = ٢١٤

وبحل هذه المعادلات آنيا عن طريق ضرب الأولى في (٥) ثم الطرح من الثانية ، نجد أن :

٤٠ $\bar{س}$ – ٥٠ $\bar{ص}$ = – ٣٣٠

$$\frac{٤٠\ \bar{س} - ١٨\ \bar{ص}}{-٣٢\ \bar{ص}} = \frac{٢١٤}{-٥٤٤} \quad ∴ \bar{ص} = \frac{٥٤٤}{٣٢} = ١٧$$

وبالتعويض في أي معادلة نحصل على $\bar{س}$

٨ $\bar{س}$ – ١٠ $\bar{ص}$ = – ٦٦ ∴ ٨ $\bar{س}$ – ١٠ × ١٧ = – ٦٦

٨ $\bar{س}$ = – ٦٦ + ١٧٠ = ١٠٤ ، $\bar{س}$ = ١٠٤ ÷ ٨ = ١٣

$\bar{س}$ = ١٣ ، $\bar{ص}$ = ١٧

(٢) استنتاج قيمة معامل الارتباط (ر) :

نفرض أن المعادلة الأولى تمثل انحدار ص / س :

$$∴ ص = \frac{٦٦}{١٠} + \frac{٨}{١٠} س$$ ، ∴ ميل خط الانحدار (ب) = $\frac{٨}{١٠}$

ولنفرض أن المعادلة الثانية تمثل انحدار س / ص :

س = $\frac{٢١٤}{٤٠}$ + $\frac{١٨}{٤٠}$ ص ∴ ميل خط الانحدار (د) = $\frac{١٨}{٤٠}$

∴ ر = $\sqrt{ب \times د}$ = $\sqrt{\frac{٨}{١٠} \times \frac{١٨}{٤٠}}$ = ٠,٦

(٣) ايجاد الانحراف المعياري للظاهرة (ص) :

∴ ميل خط الانحدار ص على س = $\frac{٨}{١٠}$ = ب

∴ ب = ر × $\frac{ع_{ص}}{ع_{س}}$ حيث ع س = $\sqrt{٩}$ = ٣

$\frac{٨}{١٠}$ = ٠,٦ × $\frac{ع_{ص}}{٣}$ ∴ عص = $\frac{٨}{١٠}$ × ٣ × $\frac{١٠}{٦}$ = ٤

مثال (٦):

بفرض أنه توفرت لديك البيانات التالية :

	الظاهرة (س)	الظاهرة (ص)
الوسط الحسابي	١٨	١٠٠
الانحراف المعياري	١٤	٢٠
معامل الارتباط بين س ، ص = + ٠,٨		

المطلوب :

(١) القيمة المتوقعة للمتغير (ص) إذا كانت س = ٧٠

(٢) القيمة المتوقعة للمتغير (س) إذا كانت ص = ٩٠

الحل :

١- للتنــبؤ بقــيمة (ص) نستخدم معادلة خط انحدار ص على س . وفي ظل المعلومات المتوفرة في هذا المثال ، نستخدم احدى صور معادلة خط انحدار ص / س التي سبق ذكرها وتتوفر فيها هذه المعلومات .

$$\therefore (ص - \bar{ص}) = ر \times \frac{ع_{ص}}{ع_{س}} (س - \bar{س})$$

$$(ص - ١٠٠) = ٠,٨ \times \frac{٢٠}{١٤} (س - ١٨)$$

∴ ص = ١,١٤ س – ٢٠,٥٢ + ١٠٠

∴ ص = ٧٩,٤٨ + ١,١٤ س

∴ القيمة المتوقعة لـــ ص عندما تكون س = ٧٠ هي :

ص = ٧٩,٤٨ + ١,١٤ × ٧٠ = ٧٩,٤٨ + ٧٩,٨ = ١٥٩,٢٨

لاحــظ أن القيمة ١٥٩,٢٨ هي القيمة المتوقعة للمتغير ص عندما تكون قيمة س = ٧٠، بمعنى أنه لو كان هناك العديد من قيم س ، كل منها تأخذ القيمة ٧٠ ، نجــد أن هــناك العديد من قيم ص المختلفة ، لكن متوسط تلك القيم (أي القيمة المتوقعة) هو ١٥٩,٢٨ وهي القيمة الواقعة على خط الانحدار .

٢- للتنــبؤ بقيمة (س) ، نستخدم معادلة خط انحدار س على ص . وفي ظل المعلومــات المــتوفرة فــي هذا المثال ، نستخدم الصورة التالية من صور معادلة خط انحدار س / ص

$$(س - \bar{س}) = ر \times \frac{ع_{س}}{ع_{ص}} (ص - \bar{ص})$$

$$(س - ١٨) = ٠,٨ \times \frac{١٤}{٢٠} (ص - ١٠٠)$$

س = ٠,٥٦ ص – ٥٦ + ١٨ ∴ س = – ٣٨ + ٠,٥٦ ص

القيمة المتوقعة لـــ س عندما تكون ص = ٩٠ هي

س = – ٣٨ + ٠,٥٦ × ٩٠ = –٣٨ + ٥٠,٤ = ١٢,٤

مثال (٧):

استخدم المعلومات التالية في ايجاد معادلة خط انحدار ص على س :

	الظاهرة (س)	الظاهرة (ص)
حجم العينة .	١٥	١٥
الوسط الحسابي .	٢٥	١٨
مجموع مربعات انحرافات القيم عن وسطها الحسابي.	١٣٦	١٢٨
مجموع حاصل ضرب انحرافات (س) ،(ص) عن أوساطهما الحسابية .	١٢٢	--

الحل :

معادلة خط انحدار ص / س هي :

$$ص = أ + ب س \qquad \text{حيث} \qquad ب = ر \times \frac{ع_ص}{ع_س}$$

$$\text{لكن } ر = \frac{مجـ (س - \bar{س})(ص - \bar{ص})}{\sqrt{مجـ (س - \bar{س})^2 \times مجـ (ص - \bar{ص})^2}} = \frac{١٢٢}{\sqrt{١٣٦ \times ١٢٨}} = ٠,٩٢٥$$

$$ع^2_ص = \frac{١}{ن - ١} \; مجـ (ص - \bar{ص})^2 \qquad ...(٣١)$$

$$= \frac{١}{١٤} \times ١٢٨ = ٩,١٤ \qquad \therefore ع_ص = \sqrt{٩,١٤} = ٣,٠٢$$

$$ع^2_س = \frac{١}{ن - ١} \; مجـ (س - \bar{س})^2 = \frac{١}{١٤} \times ١٣٦ = ٩,٧١$$

$$\therefore ع_س = \sqrt{٩,٧١} = ٣,١٢$$

$$\therefore ب = ر \times \frac{ع_ص}{ع_س}$$

$$= ٠,٩٢٥ \times \frac{٣,٠٢}{٣,١٢} = ٠,٨٩٦$$

أما أ = صَ – ب سَ

= ١٨ – ٠,٨٩٦ × ٢٥ = ١٨–٢٢,٤ = –٤,٤

∴ معادلة خط انحدار ص / س تصبح على الصورة :

ص = ٤,٤٠٠ + ٠,٨٩٦ س

(٥) الخطأ المعياري لتقدير معادلة خط الانحدار :

ذكرنا من قبل أنه من الأهداف الرئيسية لمعادلة خط الانحدار ، هو استخدامها في عملية التنبؤ بقيمة المتغير التابع عند قيمة معلومة للمتغير المستقل . وطالما أن الارتباط بين الظواهر الطبيعية غير تام بصفة عامة ، فلابد وأن تنحرف بعض القيم الفعلية للمتغير التابع عن القيم المتوقعة أو التقديرية أو النظرية ، وهي القيم التي نقدرها من معادلة خط الانحدار . وطريقة المربعات الصغرى – كما ذكرنا من قبل – تهدف أساسا إلى تصغير هذه الانحرافات والوصول بها إلى أقل قيمة ممكنة . لكن يكون من المهم والمفيد قياس مدى تشتت هذه الانحرافات ، أو بمعنى أدق قياس تشتت القيم الفعلية للمتغير التابع عن قيمها المتوقعة أي عن خط الانحدار ، وهذا القياس هو ما نعني به الخطأ المعياري standard error، ويختلف مفهوم الانحراف المعياري (الذي سبق دراسته) عن الخطأ المعياري ، فالانحراف المعياري standard deviation يقيس تشتت المفردات حول وسطها الحسابي بينما الخطأ المعياري يقيس تشتت التقديرات حول خط الانحدار .

<u>(أ) الخطأ المعياري لتقدير معادلة خط الانحدار ص / س : ص = أ + ب س</u>

لقياس تشتت القيم الفعلية (ص) حول خط الانحدار ، نستخدم العلاقة التالية :

$$خ^{٢}_{ص} = \frac{مجـ (ص - \hat{ص})^{٢}}{ن - ٢} \qquad ...(٣٢)$$

وهــذه العلاقة توضيح تباين القيم الفعلية ($\hat{ص}$) عن القيم المتوقعة ص ، والــناتجة عن معادلة خط انحدار (ص / س) ، والجذر التربيعي يعطي الخطأ المعياري للتقدير .

ومن الواضح أنه كلما كانت قيمة $خ^2_{س}$ كبيرة ، كلما دل ذلك على تباعد أو تشــتت مفــردات المتغير التابع عن خط الانحدار ،مما يعني في نفس الوقت انخفــاض قيمة معامل الارتباط بين الظاهرتين (س ، ص) والعكس صحيح ، وســوف نوضــح فــيمــا بعد العلاقة التي تربط بين الخطأ المعياري ومعامل الارتباط .

جديــد بالذكــر أن مقــام التباين في المعادلة الأخيرة (٣٢) ما هو إلا درجــات الحرية ، وهي عبارة عن حجم العينة مطروحا منها عدد المعالم التي يــتم تقديرها من العينة وهي (أ ، ب) أي أن درجات الحرية = ن – ٢ . وقد درج البعض ، بدافع السهولة إلى استخدام (ن) فقط بدلا من (ن – ٢) ، إلا أن هذا الاستخدام يتسم بعدم الدقة .

ولحســاب الخطــأ المعــياري الموضح بالمعادلة الأخيرة (٣٢) تتبع الخطوات التالية :

١- أوجد معادلة خط انحدار ص / س وهي ص = أ + ب س

٢- أوجد القيم المتوقعة أو النظرية ($\hat{ص}$) باستخدام المعادلة السابقة عند قيم س المختلفة .

٣- احسب الفرق بين القيم الفعلية ص والمتوقعة $\hat{ص}$ ثم ربع هذا الفرق $(ص - \hat{ص})^2$

٤- أوجد مجموع مربعات الفروق السابقة مجـ $(ص - \hat{ص})^2$

٥- طبق علاقة الخطأ المعياري $خ_{ص}$ التالية :

$$خ_{ص} = \sqrt{\frac{مجـ (ص - \hat{ص})^2}{ن - ٢}} \quad ...(٣٣)$$

وفي الواقع فإننا لسنا مضطرين دائما لحساب القيم النظرية ($\hat{ص}$) من معادلـــة خـــط الانحـــدار لكي نصل إلى قيمة الخطأ المعياري ، لأن هذا يشكل صـــعوبة بالغـــة في اجراء العمليات الحسابية ، ولكن يمكن أن نستعين بعلاقة أخرى أكثر بساطة لحساب الخطأ المعياري وهي :

$$خ_{ص} = \sqrt{\frac{(مجـ ص^{2} - أ\ مجـ ص - ب\ مجـ س ص)}{ن - 2}} \quad ..(34)$$

حيـــث مجـــ ص ، مجـــ ص٢ ، مجـــ س ص بيانات موجودة بالفعل وسبق استخدامها عند إيجاد معادلة خط الانحدار، كما أن أ ، ب قيم تم تقديرها .

ملحوظة :

في علم الإحصاء يفرق دائما بين الانحراف المعياري والخطأ المعياري ، فالانحـــراف المعـــياري يقصـــد به الجذر التربيعي لمتوسط مجموع مربعات انحـــرافات القـــيم عن وسطها الحسابي ، أما الخطأ المعياري فيقصد به الجذر التربيعي لمتوسط مجموع مربعات انحرافات التقديرات (التي نحصل عليها من العديـــد مـــن العيـــنات) عن المتوسط العام لهذه التقديرات (متوسط المجمع) . وبمعنى أكثر اختصارا ، الانحراف المعياري يقيس درجة تشتت المفردات حول وســـطها الحســـابي أمـــا الخطـــأ المعياري فيقيس درجة تشتت التقديرات حول متوسطها العام .

(ب) الخطأ المعياري لتقدير معادلة خط انحدار س/ص : س= جـ+ د ص

لقياس تشتت قيم المتغير التابع (س) حول خط الانحدار ، أي بعداً عن القيم التقديرية $\hat{س}$ ، نستخدم العلاقة التالية :

$$خ^{2}_{س} = \frac{مجـ (س - \hat{س})^{2}}{ن - 2} \quad ...(35)$$

حيث $\hat{س}$ هي القيم التقديرية الناتجة عن استخدام معادلة خط انحدار س /ص ، و يمكن استخدام علاقة أخرى مماثلة للعلاقة رقم (٣٤) وهى :

$$خ^2_س = \frac{مجـ س^2 - جـ مجـ س - د مجـ س ص}{ن - ٢} \quad ...(٣٦)$$

و الخطأ المعياري هو الجذر التربيعي للتباين في المعادلات (٣٥) ، (٣٦)

مثال (٨):

استخدم بيانات المثال رقم (٣) في إيجاد الخطأ المعياري لتقديرات معادلة خط انحدار ص/س

الحل :

من مثال رقم (٣) وجدنا أن معادلة خط انحدار ص/س كانت على الصورة التالية :

ص = ٢٧,١٢ + ١,٦٦ س

وكما ذكرنا من قبل ، فهناك صيغتان لحساب الخطأ المعيارى ، الأولى بالمعادلة رقم (٣٣) و الثانية بالمعادلة (٣٤) وسوف نعرض الحلين بغرض توضيح كيفية الإستخدام.

(أ) $$خ_ص = \sqrt{\frac{مجـ (ص - \hat{ص})^2}{ن - ٢}}$$

وهنا يلزم إيجاد القيم التقديرية $\hat{ص}$ باستخدام معادلة خط الانحدار ، وذلك عند جميع قيم (س) المعطاة في المثال ، و الجدول رقم (٤) يصور العمليات الحسابية اللازمة لقياس الخطأ المعياري .

جدول (٤)

س	ص	$\hat{ص}$ = ٢٧,١٢ + ١,٦٦ س	ص – $\hat{ص}$	(ص – $\hat{ص}$)^{٢}
٦	٤٠	٢٧,١٢ + ١,٦٦ × ٦ = ٣٧,٠٨	+ ٢,٩٢	٨,٥٢٦٤
١٠	٤٤	٢٧,١٢ + ١,٦٦ × ١٠ = ٤٣,٧٢	+٠,٢٨	٠,٠٧٨٤
١٢	٤٦	= ٤٧,٠٤	– ١,٠٤	١,٠٨١٦
١٤	٤٨	= ٥٠,٣٦	– ٢,٣٦	٥,٥٦٩٦
١٦	٥٢	= ٥٣,٦٨	– ١,٦٨	٢,٨٢٢٤
١٨	٥٨	= ٥٧,٠٠	+١	١,٠٠٠
٢٢	٦٠	= ٦٣,٦٤	–٣,٦٤	١٣,٢٤٩٦
٢٤	٦٨	= ٦٦,٩٦	+١,٠٤	١,٠٨١٦
٢٦	٧٤	= ٧٠,٢٨	+٣,٧٢	١٣,٨٣٨٤
٣٢	٨٠	٢٧,١٢ + ١,٦٦ × ٣٢ = ٨٠,٢٤	–٠,٤٢	٠,٠٥٧٦
			صفر	٤٧,٣٠٥٦

∴ الخطأ المعيارى خم = $\sqrt{\frac{٤٧,٣٠٥٦}{١٠-٢}} = \sqrt{\frac{٤٧,٣٠٥٦}{٨}}$

$= \sqrt{٥,٩١٣٢} = ٢,٤٣$

(ب) خم = $\sqrt{\frac{\text{مجـ ص}^{٢} - \text{أ مجـ ص} - \text{ب مجـ س ص}}{\text{ن} - ٢}}$

وهنا يلزم معرفة مجـ ص^{٢} ، مجـ س ص ، مجـ ص ، ولم تكن قد حسبت فى جدول (٢) أما أ= ٢٧,١٢ ، ب = ١,٦٦ .

جدول (٥)

س	ص	س ص	$ص^{٢}$
٦	٤٠	٢٤٠	١٦٠٠
١٠	٤٤	٤٤٠	١٩٣٦
١٢	٤٦	٥٥٢	٢١١٦
١٤	٤٨	٦٧٢	٢٣٠٤
١٦	٥٢	٨٣٢	٢٧٠٤
١٨	٥٨	١٠٤٤	٣٣٦٤
٢٢	٦٠	١٣٢٠	٣٦٠٠
٢٤	٦٨	١٦٣٢	٤٦٢٤
٢٦	٧٤	١٩٢٤	٥٤٧٦
٣٢	٨٠	٢٥٦٠	٦٤٠٠
١٨٠	٥٧٠	١١٢١٦	٣٤١٢٤

$$\therefore \text{خ}_{\text{ص}} = \sqrt{\frac{\text{٣٤١٢٤} - \text{٢٧,١٢} \times \text{٥٧٠} - \text{١,٦٦} \times \text{١١٢١٦}}{\text{١٠} - \text{٢}}}$$

$$= \sqrt{\frac{\text{٣٤١٢٤} - \text{٣٤٠٧٦,٩٦}}{\text{٨}}} = \sqrt{\frac{\text{٤٧,٠٤}}{\text{٨}}} = \sqrt{\text{٥,٨٨}} = \text{٢,٤٢}$$

وهى نفس النتيجة السابقة ولكن مع اختلاف ضئيل نتيجة للقيم التقريبية لكل من أ ، ب .

(٦) العلاقة بين الخطأ المعياري لمعادلة خط الانحدار ومعامل الارتباط:

بينا من قبل أنه إذا كانت النقط الانتشارية قريبة من خط الانحدار ، دل ذلك على وجود علاقة ارتباط قوية بين الظاهرتين (س ،ص) والعكس صحيح ،

فكلما تباعدت النقط الانتشارية عن خط الانحدار كلما دل ذلك على ضعف علاقة الارتــباط بين الظاهرتين . تشتت النقط حول خط الانحدار المستقيم عبارة عن الخطأ المعياري والذي تناولناه فى البند السابق .

مــن تلــك الملاحظة نجد أن هناك <u>علاقة عكسية</u> بين الخطأ المعياري ومعـــامل الارتباط، فكلما كان تشتت النقط حول خط الانحدار (الخطأ المعياري) قلــيلاً ، كلما كان معامل الارتباط كبيراً والعكس صحيح . ومن الممكن ترجمة هــذه النتيجة فى صورة علاقة رياضية سواء أكان خط الانحدار هو ص/س أو س/ص .

<u>أ- إذا كانت علاقة خط الانحدار هى : ص = أ + ب س :</u>

$$\text{خ}^{٢}_{\text{ص}} = \text{ع}^{٢}_{\text{ص}} \left(١ - \text{ر}^{٢} \right) \qquad \ldots (٣٧)$$

أى أن

$$\boxed{\text{خ}_{\text{ص}} = \text{ع}_{\text{ص}} \times \sqrt{١ - \text{ر}^{٢}}} \qquad \ldots (٣٨)$$

أى أنه إذا علم معامل الارتباط (ر)، يمكن استنتاج الخطأ المعياري خص وهذه صورة أخرى من صور الخطأ المعياري بجانب المعادلات رقم (٣٣،٣٤).

ومـــن الممكـــن اســـتنتاج قـــيمة معامل الارتباط بدلالة معلومية الخطأ المعياري من العلاقة (٣٨) على النحو التالى :

$$\text{ر} = \sqrt{١ - \frac{\text{خ}^{٢}_{\text{ص}}}{\text{ع}^{٢}_{\text{ص}}}} \qquad \ldots (٣٩)$$

حيث $\text{ع}^{٢}_{\text{ص}}$ في المعادلات الأخيرة هي تباين المفردات الفعلية للمتغير التابع ص ، أى أن :

$$\text{ع}^{٢}_{\text{ص}} = \frac{١}{\text{ن} - ١} \text{ مجـ } (\text{ص} - \bar{\text{ص}})^{٢} = \frac{١}{\text{ن} - ١} \left(\text{مجـ ص}^{٢} - \frac{(\text{مجـ ص})^{٢}}{\text{ن}} \right) \qquad \ldots (٤٠)$$

وعلى ذلك : $ع^2_س$: تباين مفردات ($\bar{ص}$) حول متوسطها ص .

$خ^2_س$: تباين مفردات (ص) حول خط الانحدار ص/س .

ب- إذا كانت معادلة خط الانحدار هي : س = جـ + د ص :

فإن العلاقة التي تربط بين الخطأ المعياري ومعامل الارتباط تكون على الصورة:

$$خ_س = ع_س \sqrt{1 - ر^2} \quad ...(41)$$

حيث $ع_س$: الانحراف المعياري لمفردات المتغير التابع س عن الوسط الحسابي $\bar{س}$.

$خ_س$: الخطأ المعياري لمفردات المتغير التابع عن خط الانحدار أى عن القيم $\hat{س}$.

$ر^2$: مربع معامل الارتباط ويسمي بمعامل التحديد (وسوف نتناوله تفصيلاً فيما بعد) وجذره التربيعي هو معامل الارتباط .

من ناحية أخرى يمكن استنتاج قيمة (ر) بدلالة $خ_س$:

$$ر = \sqrt{1 - \frac{خ^2_س}{ع^2_س}} \quad ... (42)$$

ويلاحظ من المعادلات (38 ، 41) أنه فى حالة الارتباط التام (طردى أو عكسي) بين الظاهرتين س ، ص أى أن : ر = $\pm$ 1 فإننا نجد :

$خ_ص = ع_ص \times$ صفر = صفر .

$خ_س = ع_س \times$ صفر = صفر .

أى أنـــه فـــى حالة الارتباط التام بين الظاهرتين ، فإن خطأ التقدير أو الخطأ المعياري يساوى الصفر لأن الارتباط التام يعنى أن جميع النقط تقع على خط الانحدار (ص/س او س/ص) وبالتالى لا يوجد تشتت لقيم المتغير التابع .

مثال (٩):

مستخدماً نتائج مثال (٨) قدر قيمة معامل الارتباط (ر) بين كمية الإنتاج (ص) وكمية السماد (س) .

الحل :

من نتائج مثال (٨) توصلنا إلى أن الخطأ المعياري لمفردات المتغير التابع ص أى خ ص = ٢,٤٣ ، $خ^{٢}_{ص}$ = ٥,٨٨ .

$$\therefore ر = \sqrt{١ - \frac{خ^{٢}_{ص}}{ع^{٢}_{ص}}}$$

من الواضح أننا نحتاج إلى معرفة قيمة $ع^{٢}_{ص}$:

$$\therefore ع^{٢}_{ص} = \frac{١}{ن-١} \text{مجـ} (ص - \bar{ص})^{٢} = \frac{١}{ن-١} (\text{مجـ ص}^{٢} - \frac{(\text{مجـ ص})^{٢}}{ن})$$

ومن الجدول (٥) نجد أن :

$$ع^{٢}_{ص} = \frac{١}{٩} (٣٤١٢٤ - \frac{(٥٧٠)^{٢}}{١٠})$$

$$= \frac{١}{٩} (٣٤١٢٤ - ٣٢٤٩٠) = \frac{١٦٣٤}{٩} = ١٨١,٥٦$$

$$\therefore ر = \sqrt{١ - \frac{٥,٨٨}{١٨١,٥٦}} = \sqrt{١ - ٠,٠٣٢} = \sqrt{٠,٩٦٨} = ٠,٩٨٤$$

وللــتأكد من صحة هذه النتيجة نقوم بحساب معامل الارتباط من بيانات المثال الأصلى وهو مثال (٣) وباستخدام الجدول (٢) :

$$ر = \frac{\text{مجـ ح}_{١}\text{ح}_{٢} - \frac{\text{مجـ ح}_{١} \times \text{مجـ ح}_{٢}}{ن}}{\sqrt{(\text{مجـ ح}^{٢}_{١} - \frac{(\text{مجـ ح}_{١})^{٢}}{ن})(\text{مجـ ح}^{٢}_{٢} - \frac{(\text{مجـ ح}_{٢})^{٢}}{ن})}}$$

$$= \frac{٨١٦ - (-٢٠ \times ٧٠ / ١٠)}{\sqrt{[٦١٦ - (٢٠)^٢/١٠][٢١٢٤ - (٧٠)^٢/١٠]}} = \frac{٨١٦ + ١٤٠}{\sqrt{(٦١٦ - ٤٠)(٢١٢٤ - ٤٩٠)}}$$

$$= \frac{٩٥٦}{\sqrt{٥٧٦ \times ١٦٩٣٤}} = \frac{٩٥٦}{٩٧٠,١٥} = ٠,٩٨٥$$

وهى نفس النتيجة التي حصلنا عليها من قبل ، ولكن مع اختلاف طفيف حوالي ٠,١ % .

مثال (١٠):

إذا كانت معادلتي خط انحدار ص/س ، س/ص على النحو التالي :

ص = ١٠ + ٠,٣ س ... (ص/س)

س = - ٤ + ١,٦ ص ... (س/ص)

فما هي قيمة الخطأ المعياري لتقديرات معادلة خط انحدار س على ص ، إذا علمت أن تباين مفردات الظاهرة (س) = ٤ .

الحل :

من معادلتي خط الانحدار يمكن استنتاج قيمة معامل الارتباط .

$$\therefore ر^٢ = ب \times د = ٠,٣ \times ١,٦ = ٠,٤٨$$

$$\therefore خ_{س} = ع_{س} \sqrt{١ - ر^٢} \quad \text{حيث} \quad ع_{س} = \sqrt{٤} = ٢$$

$$\therefore خ_{س} = ٢ \times \sqrt{١ - ٠,٤٨} = ٢ \times \sqrt{٠,٥٢} = ١,٤٤$$

(٧) معامل التحديد ومعامل الارتباط:

في بداية الحديث عن موضوع الانحدار ، ذكرنا أن من أهداف دراسة هذا الموضوع هو معرفة درجة تأثير المتغير المستقل (س مثلا) على المتغير التابع (ص) ، فقد يكون مهما معرفة درجة تأثير السماد (متغير مستقل) على حجم الإنتاج (متغير تابع) . نعلم ان حجم الإنتاج يتوقف ويعتمد على عدة

عوامـــل مـــنها السماد ، كمية المياه ، درجة الحرارة ، درجة الرطوبة ، مهارة العـــامل ، نوع التربة ... إلخ . فإذا قسمنا مجموعة هذه العوامل إلى قسمين : قســـم يضـــم السماد (وهو المتغير المستقل الذي نريد معرفة درجة تأثيره على المتغـــير الـــتابع) وقســـم ثانى يضم باقى العوامل الأخرى ، ويطلق عليها اسم مجموعة العوامل العشوائية ، نجد أن التغير الذي يحدث فى الإنتاج يرجع إلى : التغير الذي يحدث فى المتغير المستقل (السماد) والتغير الذي يحدث فى العوامل العشوائية .

يطلـــق علـــى التغـــير الذي يحدث فى الإنتاج ، أى فى المتغير التابع ، بالتغير الكلى Total Variation أما التغير الذي يحدث فى المتغير المستقل ، فـــيطلق عليه اسم التغير المفسر Explained Variation ، بينما التغير الذي يحـــدث فـــى المتغـــيرات او العوامـــل العشـــوائية يسمي بالتغير غير المفسر Unexplained Variation .أى أن التغيرات الكلية فى الإنتاج (متغير تابع) زيـــادة أو نقص ، تتكون من شقين : شق يمكن تفسيره على أنه ناتج عن وجود السماد (متغير مستقل) أثر فى كمية الإنتاج ، وشق آخر لا يمكن تفسيره بعامل محدد ، ولكنه بصفة عامة يرجع إلى خليط من عوامل عشوائية لا يمكن التحكم فيها (وهو افتراض فى ظل هذه التجربة وإن كان يمكن إخراج بعض المتغيرات المســـتقلة مـــن هذه العوامل العشوائية مثل كمية المياه ودرجة الحرارة كعوامل مؤثـــرة فـــى الإنتاج بجانب السماد وهذا ينقلنا إلى موضوع آخر وهو الانحدار المتعدد) .

∴ التغـــير الكلـــى فـــى المتغير التابع = التغير المفسر (بسبب وجود المتغير المستقل) + التغير غير المفسر (عوامل عشوائية) .

جدير بالذكر أنه فى علم الإحصاء نفرق بين اصطلاحي التغير والتباين Variation & Variance فالتغـــير يقصد به مجموع مربعات انحرافات القيم عـــن وســـطها الحسابي (بالرموز : التغير = مجـــ (ص – $\bar{ص}$)^2 ، أما التباين

فيقصـد بـه متوسـط مجموع مربعات انحرافات القيم عن وسطها الحسابي . [بالـرموز : التباين = مجـ (ص − $\bar{ص}$)^2 / (ن−١)] . فى ضوء هذه التفرقة نجد أن :

١− التغير الكلى فى المتغير التابع (ص) = مجـ (ص − $\bar{ص}$)^2 ... (٤٣)

$$= \text{مجـ ص}^2 - \frac{(\text{مجـ ص})^2}{\text{ن}} \quad ... (٤٤)$$

٢−التغـير المفسـر الناشـئ عن وجود المتغير المستقل (س) أى الناشئ عن استخدام معادلة خط

الانحدار ص/س والتي تعطى القيم التقديرية $\hat{ص}$ = مجـ ($\hat{ص}$ − $\bar{ص}$)^2

= ب مجـ (س − $\bar{س}$) (ص − $\bar{ص}$) ... (٤٥)

$$= \text{ب} \left(\text{مجـ س ص} - \frac{\text{مجـ س} \times \text{مجـ ص}}{\text{ن}} \right) \quad ... (٤٦)$$

٣− التغـير غير المفسر والذي يرجع إلى العوامل العشوائية ، وهو عبارة عن الفـرق بين القيم الفعلية ($\hat{ص}$) والقيم التقديرية ص والتي حصلنا عليها من خط الانحدار = مجـ (ص − $\hat{ص}$)^2

∴التغير الكلى فى (ص)=التغير المفسر بسبب وجود س+ التغيّر غير المفسر (الخطأ العشوائي)

مجـ (ص − $\bar{ص}$)^2 = مجـ ($\hat{ص}$ − $\bar{ص}$)^2 + مجـ (ص − $\hat{ص}$)^2

يلاحظ أنه عند ثبات حجم التغير الكلى ، فإن أى نقص فى حجم الخطأ العشـوائي ، يعني زيادة فى حجم التغير المفسر الناشئ عن استخدام علاقة خط الانحـدار ، مما يعني فى نفس الوقت زيادة تأثير المتغير المستقل على المتغير الـتابع ، ومما يعني أيضاً زيادة قيمة معامل الارتباط بين المتغيرين (س،ص) والعكس صحيح .

ويعــرف معامل التحديد Coefficient of Determination على أنه نسبة التغير المفسر إلى التغير الكلى ويرمز له بالرمز ر^{٢} أى أن :

$$\text{معامل التحديد } ر^{٢} = \frac{\text{التغير المفسر}}{\text{التغير الكلى}}$$

$$= \frac{\text{مجـ}(\hat{ص} - \bar{ص})^{٢}}{\text{مجـ}(ص - \bar{ص})^{٢}} \quad ...(٤٧)$$

$$= \frac{ب\ \text{مجـ}(س - \bar{س})(ص - \bar{ص})}{\text{مجـ}(ص - \bar{ص})^{٢}} \quad ...(٤٨)$$

$$= \frac{ب\ (\text{مجـ}\ س\ ص - \text{مجـ}\ س \times \text{مجـ}\ ص / ن)}{\text{مجـ}\ ص^{٢} - (\text{مجـ}\ ص)^{٢} / ن} \quad ...(٤٩)$$

مــن ناحية أخرى ، فإن معامل التحديد يبين نسبة تأثير المتغير المستقل (س) علــى المتغــير الــتابع (ص) أما الجذر التربيعي لمعامل التحديد فيساوي معامل الارتباط المعروف (ر).

وإذا كــان المقــدار (ر^{٢}) يســمي بمعامل التحديد فإن المقدار (١- ر^{٢}) يسمي بمعامل عدم التحديد ، وبديهي فإنه يعبر عن نسبة تأثير العوامل العشوائية علــى المتغــير التابع . وقيمة معامل التحديد (ر^{٢}) تتراوح بين الصفر والواحد الصحيح حيث نجد أن :

١- إذا كانت ر^{٢} = صفر ، فهذا يعنى أن بسط معامل التحديد يساوى الصفر أى ان التغير المفسر= صفر أى أن التغيرات الكلية فى التابع (ص) ترجع كلها إلى العوامل العشوائية أى أن نسبة تأثير المتغير المستقل س = صفر .

٢- إذا كانت ر^{٢} = ١ ، فهذا يعني أن التغير الكلى فى التابع (ص) ترجع كلها إلى وجود المتغير المستقل (س) وأن التغيرات العشوائية ذات تأثير منعدم .

٣- أما إذا كانت ر^{٢} تقع بين الصفر والواحد الصحيح ، فهذا يعني أن جزءاً من التغير الكلى فى التابع (ص) يرجع إلى وجود المتغير المستقل (س) وجزءاً

آخــر يرجع إلى العوامل العشوائية ، وبالطبع كلما زادت قيمة $ر^2$ واقتربت من الواحد الصحيح كان هذا دليلاً على تعاظم تأثير (س) على المتغير التابع والعكس صحيح .

مثال (١١) :

استخدم بيانات مثال رقم (٣) فى إيجاد قيمة معامل التحديد :

الحل :

$$ر^2 = \frac{ب\,(مجـ\,س\,ص - مجـ\,س \times مجـ\,ص\,/\,ن)}{مجـ\,ص^2 - (مجـ\,ص)^2\,/\,ن} \quad \text{(الطريقة المباشرة)}$$

أو

$$= \frac{ب\,(مجـ\,ح_1\,ح_2 - مجـ\,ح_1 \times مجـ\,ح_2\,/\,ن)}{مجـ\,ح^2_2 - (مجـ\,ح_2)^2\,/\,ن} \quad \text{(الطريقة المختصرة)}$$

بما أن معادلة خط الانحدار فى مثال (٣) كانت على الصورة : ص = ٢٧,١٢ + ١,٦٦ س أى أن ميل خط الانحدار ب = ١,٦٦ . من جدول (٥) نجد أن :

$$ر^2 = \frac{١,٦٦\,(١١٢١٦ - ١٨٠ \times ٥٧٠/١٠)}{٣٤١٢٤ - (٥٧٠)^2\,/\,١٠}$$

$$= \frac{١,٦٦ \times ٩٥٦}{١٦٣٤} = \frac{١٥٨٦,٩٦}{١٦٣٤} = ٠,٩٧١٢$$

ومن جدول (٢) نجد أن :

$$ر^2 = \frac{١,٦٦[٨١٦ - (-٢٠ \times ٧٠)/١٠]}{٢١٢٤ - (٧٠)^2\,/\,١٠} = \frac{١,٦٦ \times ٩٥٦}{١٦٣٤} = ٠,٩٧١٢$$

أى أن هناك تطابق تام بين نتائج الطريقتين المباشرة والمختصرة .

∴ نســبة تأثير المتغير المستقل (السماد فى هذا المثال) على كمية الإنتاج تبلغ ٩٧,١٢% ، والباقى ٢,٨٨% ترجع إلى مجموعة عوامل عشوائية . أما الجذر التربيعي لمعامل التحديد = $\sqrt{٠,٩٧١٢}$ = ٠,٩٨٥ فهو يمثل قيمة معامل الارتباط بين الظاهرتين (س،ص) وهى نفس النتيجة التى توصلنا إليها فى مثال (٩) .

ملحوظة :

يمكــن اســتخدام الصيغة رقم (٤٧) لحساب معامل التحديد لكنها تمثل مجهود حسابي ضخم ومرهق .

(٨) معادلة خط الانحدار من بيانات مبوبة :

لاحظنا فى الأمثلة التوضيحية السابقة أننا نتعامل مع عينات ذات أحجام صــغيرة (ن أقل من ٣٠) ، وذلك بغرض توضيح كيفية استخدام المقاييس التى قدمــناها ، لكن صغر حجم العينة يجعلنا تحت رحمة اى خطأ يحدث بالمصادفة عند رصد قيمة أحد المتغيرات ، ناهيك عن أن أى خطأ يحدث يترتب عليه خطأ كبــير فى النتائج التى نتوصل إليها لذلك – وكما أكدنا من قبل – يجب ألا يقل عــدد الحالات التى نتعامل معها عن ٣٠ حالة (أى أن ن أكبر من ٣٠ وتسمي عينات كبيرة) حتى تكون هناك فرصة لتعادل أثر الحالات الشاذة مع بعضها .

ولكــن إذا زاد عدد الحالات كثيراً ، فلاشك ان العمل الحسابي سيكون مضاعفاً ومرهقاً إذا ما استخدمنا أسلوب العمل السابق فى حالة البيانات المفردة ، لــذا علينا بتحويل هذه البيانات الكبيرة الحجم من حالة بيانات مفردة إلى حالة بيانات مبوبة ، كما وضحنا ذلك – فى مرحلة سابقة – عند الحديث عن معامل الارتباط فى حالة البيانات المبوبة .

والآن نتــناول كيفية تقدير معادلتى خط الانحدار ص/س ، س/ص من جدول تكرارى مزدوج وخير وسيلة للتوضيح هى مثال عملى .

مثال (١٣) :

الجـــدول التالي يوضح توزيع عينة من الأزواج (س) والزوجات (ص) عند بداية سن الزواج لكل منهم . والمطلوب :

١- تقدير عمر الزوج عندما يكون عمر الزوجة ٣٠ سنة .

٢- تقدير عمر الزوجة عندما يكون عمر الزوج ٤٥ سنة .

٣- قياس معامل الارتباط بين عمر الزوج وعمر الزوجة .

جدول (٦)

مراكز فئات س	مراكز فئات ص	١٧	١٩	٢١	٢٣	٢٥	المجموع
س	س \ ص	-١٦	-١٨	-٢٠	-٢٢	-٢٤	
٢٠	-١٨	٢			٣		٥
٢٤	-٢٢	٣	٥			٢	١٠
٢٨	-٢٦	١	٤	١٢	٥		٢٢
٣٢	-٣٠		٦	٨		٣	١٧
٣٦	-٣٤		٣	٥	٤		١٢
٤٠	-٣٨					٤	٤
	المجموع	٦	١٨	٢٥	١٢	٩	٧٠

الحل :

١- لـــتقدير عمـــر الـــزوج (س) عندما يكون عمر الزوجة ص = ٣٠ سنة ، يقتضي إيجاد معادلة خط انحدار س على ص ، حيث س = جـــ + د ص . نذكـــر انه فى حالة البيانات المفردة، كنا نجد أن كل قيمة من (س) تناظرها قيمة من (ص) ، غير أن الأمر يختلف فى حالة البيانات المبوبة ، حيث نجد أن كل فئة من (س) يناظرها العديد من التكرارات الواقعة أمام فئات مختلفة

من (ص) والعكس ، كل فئة من (ص) يناظرها عدد من التكرارات تقع أمام فئات مختلفة من (س) .

ولإيجـــاد معادلـــة خـــط انحدار س على ص من جدول مزدوج ، نقوم بحساب متوسطات قيم (س) المناظرة لكل فئة من فئات (ص) المختلفة ، أى انه فى معادلة خط انحدار س / ص :

س = جـــ + د ص ، نعتبر المتغير المستقل (ص) وكأنه مراكز فئات (ص) أما المتغـــير التابع (س) يؤخذ على أنه متوسط قيم س المناظرة لمراكز فئات (ص) السابقة .

مركز الفئة الأولى للمتغير ص = ١٧

∴ متوسط قيم (س) المقابلة لهذا المركز عبارة عن متوسط حاصل ضرب التكرار المشترك $ك_{٢١}$ (والتي تقع أسفل العمود الأول) فى مراكز فئات (س) .

∴ متوسط قيم س = $\frac{٢ \times ٢٠ + ٣ \times ٢٤ + ١ \times ٢٨}{٢ + ٣ + ١} = \frac{١٤٠}{٦} = ٢٣,٣٣$

مركز الفئة الثانية للمتغير ص = ١٩

∴ متوسط قيم س المقابلة لهذا المركز =

$$\frac{٥ \times ٢٤ + ٤ \times ٢٨ + ٦ \times ٣٢ + ٣ \times ٣٦}{٥ + ٤ + ٦ + ٣} = \frac{٥٣٢}{١٨} = ٢٩,٥٥$$

مركز الفئة الثالثة للمتغير (ص) = ٢١

متوسط قيم س المناظرة = $\frac{١٢ \times ٢٨ + ٨ \times ٣٢ + ٥ \times ٣٦}{١٢ + ٨ + ٥} = \frac{٧٧٢}{٢٥} = ٣٠,٨٨$

مركز الفئة الرابعة للمتغير (ص) = ٢٣

متوسط قيم س المناظرة = $\frac{٣ \times ٢٠ + ٥ \times ٢٨ + ٤ \times ٣٦}{٣ + ٥ + ٤} = \frac{٣٤٤}{١٢} = ٢٨,٦٧$

مركز الفئة الأخيرة للمتغير (ص) = ٢٥

$$\text{متوسط قيم س المناظرة} = \frac{٢ \times ٢٤ + ٣٢ \times ٣ + ٤٠ \times ٤}{٢ + ٣ + ٤} = \frac{٣٠٤}{٩} = ٣٣,٧٨$$

وبهـــذه الطريقة نكون قد كونا عمودين : الأول عبارة عن مراكز فئات ص (وهو المتغير المستقل فى معادلة خط انحدار س/ص) والعمود الثانى عبارة عـــن متوســـطات قـــيم س (وهو المتغير التابع هنا) ومن ثم تسهل عملية إيجاد معادلة خط الانحدار س = جـ + د ص .

جدول (٧)

س	ص	س ص	$ص^٢$
٢٣,٣٣	١٧	٣٩٦,٦١	٢٨٩
٢٩,٥٥	١٩	٥٦١,٤٥	٣٦١
٣٠,٨٨	٢١	٦٤٨,٤٨	٤٤١
٢٨,٦٧	٢٣	٦٥٩,٤١	٥٢٩
٣٣,٧٨	٢٥	٨٤٤,٥٠	٦٢٥
١٤٦,٢١	١٠٥	٣١١٠,٤٥	٢٢٤٥

س = جـ + د ص

$$د = \frac{\text{مجـ س ص} - \text{مجـ س} \times \text{مجـ ص} / ن}{\text{مجـ ص}^٢ - (\text{مجـ ص})^٢ / ن} = \frac{٣١١٠,٤٥ - \frac{١٤٦,٢١ \times ١٠٥}{٥}}{٢٢٤٥ - \frac{(١٠٥)^٢}{٥}}$$

$$= \frac{٣١١٠,٤٥ - ٣٠٧٠,٤١}{٢٢٤٥ - ٢٢٠٥} = \frac{٤٠,٠٤٥}{٤٠} = ١,٠٠١$$

$$جـ = \bar{س} - د \bar{ص} = \frac{١٤٦,٢١}{٥} - ١,٠٠١ \times \frac{١٠٥}{٥}$$

$$= ٢٩,٢٤ - ٢١,٠٢ = ٨,٢٢$$

$$\therefore س = ٨,٢٢ + ١,٠٠١ \text{ ص}$$

وعلـــى ذلك يمكن تقدير عمر الزوج (س) عندما يكون عمر الزوجة ص = ٣٠ سنة على النحو التالي : س = ٨,٢٢ + ١,٠٠١ × ٣٠ = ٨,٢٢ + ٣٠,٠٣ = ٣٨,٢٥ سنة .

٢- لـــتقدير عمـــر الـــزوجة (ص) عندما يكون عمر الزوج س = ٤٥ سنة ، يقتضي إيجاد معادلة خط انحدار ص على س ، (ص = أ + ب س) وبنفس الطـــريقة السابقة نعتبر (س) هنا وكأنها مراكز فئات المتغير المستقل (س) أما (ص) وهى المتغير التابع فتؤخذ على أنها متوسطات قيم (ص) المناظرة لمراكز فئات س السابقة .

مركز الفئة الأولى للمتغير س = ٢٠

$$\text{متوسط قيم ص المناظر له} = \frac{٢ \times ١٧ + ٣ \times ٢٣}{٢ + ٣} = \frac{١٠٣}{٥} = ٢٠,٦$$

مركز الفئة الثانية للمتغير س = ٢٤

$$\text{متوسط قيم ص المناظرة} = \frac{٣ \times ١٧ + ٥ \times ١٩ + ٢ \times ٢٥}{٣ + ٥ + ٢} = \frac{١٩٦}{١٠} = ١٩,٦$$

مركز الفئة الثالثة للمتغير س = ٢٨

$$\text{متوسط قيم ص المناظرة} = \frac{١ \times ١٧ + ٤ \times ١٩ + ١٢ \times ٢١ + ٥ \times ٢٣}{١ + ٤ + ١٢ + ٥} = \frac{٤٦٠}{٢٢} = ٢٠,٩$$

مركز الفئة الرابعة للمتغير س = ٣٢

$$\text{متوسط قيم ص المناظرة} = \frac{٦ \times ١٩ + ٨ \times ٢١ + ٣ \times ٢٥}{٦ + ٨ + ٣} = \frac{٣٥٧}{١٧} = ٢١$$

مركز الفئة الخامسة للمتغير س = ٣٦

$$\text{متوسط قيم ص المناظرة} = \frac{٣ \times ١٩ + ٥ \times ٢١ + ٤ \times ٢٣}{٣ + ٥ + ٤} = \frac{٢٥٤}{١٢} = ٢١,١٧$$

مركز الفئة الأخيرة = ٤٠

$$\text{متوسط قيم ص المناظرة} = \frac{٤ \times ٢٥}{٤} = ٢٥$$

ومن الجدول التالى يمكن تقدير معادلة خط انحدار ص على س :

جدول (٨)

س	ص	س ص	$س^{٢}$
٢٠	٢٠,٦	٤١٢	٤٠٠
٢٤	١٩,٦	٤٧٠,٤	٥٧٦
٢٨	٢٠,٩	٥٨٥,٢	٧٨٤
٣٢	٢١	٦٧٢	١٠٢٤
٣٦	٢١,١٧	٧٦٢,١٢	١٢٩٦
٤٠	٢٥	١٠٠٠	١٦٠٠
١٨٠	١٢٨,٢٧	٣٩٠١,٧٢	٥٦٨٠

ص = أ + ب س

$$\therefore ب = \frac{\text{مجـ س ص} - \text{مجـ س} \times \text{مجـ ص} / ن}{\text{مجـ } س^{٢} - (\text{مجـ س})^{٢} / ن}$$

$$= \frac{٣٩٠١,٧٢ - ١٨٠ \times ١٢٨,٢٧ / ٦}{٥٦٨٠ - (١٨٠)^{٢} / ٦} = ٠,١٩١٥$$

$$أ = \bar{ص} - ب \bar{س} = \frac{١٢٨,٢٧}{٦} - ٠,١٩١٥ \times \frac{١٨٠}{٦}$$

$$= ٢١,٣٨ - ٥,٧٥ = ١٥,٦٣$$

$$\therefore ص = ١٥,٦٣ + ٠,١٩١٥ \text{ س}$$

وعلى ذلك فعمر الزوجة (ص) عندما يكون عمر الزوج (س) ٤٥ سنة من هذه البيانات هو : ص = ١٥,٦٣ + ٠,١٩١٥ × ٤٥ = ١٥,٦٣ + ٨,٦٢ = ٢٤,٢٥ سنة

٣- قياس معامل الارتباط بين عمر الزوج وعمر الزوجة :

مـــن الواضح أنه من الأسهل استخدام العلاقة بين ميلى خطى الانحدار فـــى اســـتنتاج قيمة معامل الارتباط ، بدلاً من إيجاد معامل الارتباط من جدول مزدوج بالطريقة التقليدية .

∴ ص = ١٥,٦٣ + ٠,١٩١٥ س .. (انحدار ص/س)

س = ٨,٢٢ + ١,٠٠١ ص .. (انحدار س/ص)

∴ ر = $\sqrt{ب \times د}$

∴ = $\sqrt{٠,١٩١٥ \times ١,٠٠١}$ = ٠,٤٤

(٩) الاستدلال الإحصائي عن معالم خط الإنحدار

بيـــنا في مرحلة سابقة أن التقديرات وإختبارات الفروض الإحصائية ، يشكلا معا الأدوات الأساسية للاستدلال أو الاستنتاج الاحصائي . نظرية التقدير تـــتكون مـــن جزئيـــن : التقدير بنقطة والتقدير بفترة ثقة . اختبارات الفروض الاحصائية تتكون من : اختبارات معلمية واختبارات لا معلمية . وسوف نركز فـــي هذا الفصل على التقدير بفترة ثقة وعلى الاختبارات المعلمية لما لهما من مجال عملي واسع في تحليل الانحدار .

أولا : التقدير بفترة ثقة :

بفرض أنه من عينة عشوائية توصلنا إلى معادلة خط الانحدار التالية : ص = ٥ + ٣ س ، أي أن : أ = ٥ ، ب = ٣ وتسمى تلك التقديرات بالتقدير بنقطة Point Estimate للقيم الحقيقة في المجتمع γ ، β والسؤال هنا : إلى أي مـــدى يمكـــن الاعتماد على تلك التقديرات أ=٥ ، ب= ٣ في عملية التنبؤ

بالمتغير التابع ص؟ من المعلوم أنه وبسبب استخدام أسلوب العينة ، ستختلف التقديرات أ ، ب من عينة إلى أخرى ، كما أنها ستختلف عن القيم الحقيقية γ ، β . لكن بتعدد العينات التي يمكن سحبها من المجتمع ، سنجد أن متوسط تلك التقديرات (أي القيمة المتوقعة لها) ، لابد وأن تساوي القيمة الحقيقية في المجتمع ، أي توقع (أ) = γ ، توقع (ب)= β . دعنا نعيد صياغة السؤال السابق بطريقة أخرى : إلى أى مدى تقترب التقديرات أ ، ب من القيم الحقيقية المجهولة γ ، β ؟ للإجابة على ذلك ، فإننا نحتاج إلى قياس الخطأ المعياري لتلك التقديرات حتى يمكننا بناء فترة ثقة للمعالم الحقيقة γ ، β . لكن كيف تنشأ فترة الثقة؟ هذا يتوقف على معرفة التوزيع الاحتمالي للتقديرات أ ، ب . تحت شرط خضوع الحد العشوائي ψ للتوزيع الطبيعي بمتوسط وبتباين على الصورة :

(١) توقع (أ) = γ

$$\text{تباين (أ)} = \sigma^2(\gamma) = \frac{\text{مجـ س}^2}{\text{ن مجـ (س} - \bar{\text{س}})^2} \times \sigma^2$$

أي أن : أ ~ م [γ ، $\sigma^2(\gamma)$] ، م : تشير إلى التوزيع الطبيعي ، σ^2 هو تباين الحد العشوائي ψ وهو غالبا مجهول القيمة .

(٢) توقع (ب) = β

$$\text{تباين (ب)} = \sigma^2\text{(ب)} = \frac{1}{\text{مجـ (س} - \bar{\text{س}})^2} \times \sigma^2$$

أي أن : ب ~ م [β ، σ^2 (ب)]

وسوف نقتصر على الاستدلال الإحصائي الخاص بالمعلمة (β) فقط لما لها من أهمية خاصة في تحليل الانحدار أما الاستدلال حول (γ) فنادراً ما يعول عليه .

بتحويل ب إلى قيمة معيارية ي ، حيث :

$$\text{ي} = \frac{\text{ب} - \beta}{\sigma\text{(ب)}}$$

نجد أن ي تتبع توزيع طبيعي ومن ثم يمكن اجراء اختبارات الفروض الإحصائية وإنشاء فترات ثقة للمعلمة β . لكن هذا يتوقف

علـــى معلومـــية $6^{٢}$ (تبايـــن الحد العشوائي Ψ) وهو قيمة مجهولة ، لذا فإننا نستبدلها بتقدير غير متحيز من عينة عشوائية هو $\text{خ}^{٢}_{\text{ص}}$ وفي ظل استخدام عينة صغيرة الحجم ، يتحول المتغير ي إلى المتغير ت ، حيث :

$$\text{ت} = \frac{\text{ب} - \beta}{\text{ع (ب)}} \quad \text{حيث} \quad \text{ع}^{٢}\text{(ب)} = \frac{\text{خ}^{٢}_{\text{ص}}}{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}} \quad ، \quad \text{خ}^{٢}_{\text{ص}} = \frac{\text{مجـ}(\text{ص} - \bar{\text{ص}})^{٢}}{\text{ن} - ٢}$$

ع (ب) : الخطأ المعياري للتقدير (ب)

$\text{خ}_{\text{ص}}$: الخطأ المعياري لتقديرات معادلة خط الانحدار .

والمتغـــير ت هـــو متغير يتبع توزيع ت بدرجات حرية (ن − ٢) ، وعلـــى ذلـــك يمكن إنشاء فترة ثقة للمعلمة β عند درجة ثقة (١− α)% على النحو التالي :

$$\therefore \beta = \text{ب} \pm \text{ت}_{(\text{ن}-٢ \text{ ، } \alpha / ٢)} \times \text{ع (ب)} \qquad ...(٥٠)$$

مثال (١٣):

مستخدما بيانات الأمثلة (٣ ، ٨) قدر β بفترة ثقة ٩٥%

الحل :

من مثال (٣) توصلنا إلى : ص = ٢٧,١٢ + ١,٦٦ س ، ن = ١٠

ومن مثال (٨) توصلنا إلى : $\text{خ}_{\text{ص}}$ = ٢,٤٣ ، وأما الكمية :

$$\text{مجـ}(\text{س} - \bar{\text{س}})^{٢} = \text{مجـ س}^{٢} - (\text{مجـ س})^{٢}/\text{ن} = \text{مجـ ح}^{٢}_{١} - (\text{مجـ ح}_{١})^{٢}/\text{ن}$$

$$= ٦١٦ - \frac{(-٢٠)^{٢}}{١٠} = ٦١٦ - ٤٠ = ٥٧٦$$

$$\therefore \beta = \text{ب} \pm \text{ت}_{(\text{ن}-٢ \text{ ، } \alpha / ٢)} \times \text{ع (ب)} = ١,٦٦ \pm \text{ت}_{(٨ \text{ ، } ٠,٠٢٥)} \times \frac{٢,٤٣}{\sqrt{٥٧٦}}$$

$$= ١,٦٦ \pm ٢,٣٠٦ \times \frac{٢,٤٣}{٢٤} = ١,٦٦ \pm ٠,٢٣ = ١,٤٣ \text{ ، } ١,٨٩$$

∴ عــند درجة ثقة ٩٥% ، نجد أنه في المدى الطويل ، أنه في كل ١٠٠ فترة ثقة للمعلمة β ، نجد أن هناك ٩٥ فترة تحتوي كل منها على المعلمة الحقيقية β ، ومــن الخطــأ القول بان هناك احتمال قدرة ٠,٩٥ بأن فترة الثقة (١,٤٣ إلى ١,٨٩) تحتوى على المعلمة الحقيقية β .

ثانياً : اختبارات الفروض الإحصائية :

تهــتم اختبارات الفروض الإحصائية ببناء قاعدة أو إجراءات ، لاتخاذ قــرار بشــأن بقــبول أو رفــض الفرض العدمى ، وهناك أسلوبين لأداء تلك الاختبارات :

أ – فترات الثقة   ب– اختبارات المعنوية

أ – اختبارات الفروض : أسلوب فترات الثقة :

إذا وقعت قيمة المعلمة β المراد اختبارها فى ظل الفرض العدمى داخل فترة الثقــة (١- ∝)% فإننا نقبل الفرض العدمى ، وإذا وقعت تلك القيمة خارج حــدي الــثقة ، فإننا نرفض الفرض العدمى ، فمثلاً : إذا كانت قيمة β المراد اختبارها هي β = ٠,٣ (أى أن الفرض العدمى هو β = ٠,٣) وكانت فترة الثقة للمعلمة β على الصورة β = ٠,٤٣ ، ٠,٥٩ فمن الملاحظ وقوع قيمة β المراد اختــبارها وهى ٠,٣ خارج حدي الثقة ، مما يعني رفض الفرض العدمى ، لأن احتمال مشاهدة قيمة للمعلمة β = ٠,٣ هو احتمال ضئيل جداً قدره (∝÷٢) % أمــا إذا كانــت القــيمة المراد اختبارها للمعلمة β هى الصفر (أى أن الفرض العدمى هو β = صفر) فإننا نبحث عن وجود أو عدم وجود الصفر داخل حدي الــثقة ، فــإن وقــع الصفر داخل حدي الثقة لكان ذلك دليلاً على قبول الفرض العدمى .

ب- اختبارات الفروض : أسلوب اختبار المعنوية :

الأسـلوب الـبديل لفترة الثقة عند اختبار قيمة المعلمة β ، هو أسلوب اختبار المعنوية للمعلمة β . الفكرة الأساسية وراء اختبار المعنوية ، تعتمد على وسيلة الاختبار الإحصائي والتوزيع الاحتمالى لهذه الوسيلة ، ويتم الوصول إلى قـرار بقـبول أو رفـض الفـرض العدمى ، بعد مقارنة القيمة العددية لوسيلة الاختبار مع قيمة جدولية تتبع نفس التوزيع الاحتمالى لوسيلة الاختبار .

بينا من قبل أن القيمة المعيارية للإحصاء (ب) كانت على الصورة :

$$ت = \frac{(ب - \beta)}{ع\,(ب)} = \frac{(ب - \beta)\sqrt{مجـ\,(س - \bar{س})^{2}}}{خ_{ص}} \quad \dots (51)$$

ت هـنـا تعتبر وسلة الاختبار الإحصائي ، وهى متغير عشوائي يتبع توزيـع ت بدرجـات حرية (ن-2). عند مقارنة القيمة العددية لوسيلة الاختبار (ت) مـع قـيمة جدولـيه مستخرجة من توزيع ت عند درجات حرية (ن-2) ومستوى معنوية (α / 2) نصل إلى قرار بقبول أو رفض الفرض العدمى .

مثال (14):

مسـتخدماً بـيانات مـثال (3) ، المطلوب اختبار صلاحية معادلة خط الانحدار للاستخدام في علمية التنبؤ عند مستوى معنوية 5% ، مستخدماً في ذلك أسلوبي فترة الثقة واختبار المعنوية.

الحل :

اختبار صلاحية المعادلة للاستخدام تعنى إجراء اختبار β = صفر .

أ - اختبار β = صفر باستخدام فترة الثقة .

خطوات الاختبار :

1- الفرض العدمى : β = صفر . (لا توجد علاقة حقيقية بين س ، ص)

٢- الفرض البديل : $\beta \neq$ صفر (توجد علاقة حقيقية بين س ، ص)

٣- فترة الثقة للمعلمة β سبق الحصول عليها فى مثال (١٣) وكانت على الصورة : β =١,٤٣ إلى ١,٨٩ .

٤- وحيث أن قيمة المعلمة المراد اختبارها هى β = صفر ، وحيث أن الصفر لا يقع داخل فترة الثقة فإن القرار هو :

٥- القرار : رفض الفرض العدمى وقبول الفرض البديل

∴ $\beta \neq$ صفر ،أى هناك علاقة ارتباط حقيقية بين المتغيرين (س ،ص)

<u>ب- اختبار β = صفر باستخدام اختبار المعنوية (اختبار ت) :</u>

خطوات الاختبار :

١- الفرض العدمى : β = صفر .

٢- الفرض البديل : $\beta \neq$ صفر

٣- $$\text{ت} = \frac{(\text{ب} - \beta)}{\text{ع}(\text{ب})} = \frac{(\text{ب} - \text{صفر})}{\text{خ}_{\text{ص}} / \sqrt{\text{مجـ س}^{2} - (\text{مجـ س})^{2} / \text{ن}}} = \frac{\text{١,٦٦}}{\text{٢,٤٣} \div \text{٢٤}} = \text{١٦,٣٩٥}$$

٤- ت الجدولية : ت (ن - ٢ ، α / ٢) = ت (٨ ، ٠,٠٢٥)= ٢,٣٠٦

٥- المقارنة : ت المحسوبة أكبر من ت الجدولية .

٦- القرار : رفض الفرض العدمى ومن ثم قبول الفرض البديل

∴ $\beta \neq$ صفر وهو نفس القرار الذي توصلنا إليه باستخدام أسلوب فترة الثقة .

(١٠) تحليل التباين وتحليل الانحدار :

يمكن استخدام أسلوب تحليل التباين في اختبار مدى وجود علاقة حقيقية بين المتغير التابع (ص) والمتغير المستقل (س) ، أي في اختبار β = صفر ، وذلك بتجزئة الاختلافات الكلية في المتغير التابع (ص) إلى مركبتين : الأولى ترجع إلى اختلافات ناتجة عن العلاقة الانحدارية بين (ص ، س) أي اختلافات تفسر بسبب وجود المتغير المستقل (س) ، والثانية ترجع إلى

اخـــتلافات ناتجة عن مجموعة من المتغيرات العشوائية ، أي اختلافات لا يمكن ارجاعها أو تفسيرها إلى سبب معين ، أي اختلافات غير مفسرة ، وقد بينا عند الحديث عن معامل التحديد ما يلي :

التغير الكلي في المتغير التابع (ص) = مجموع المربعات الكلى (م.م.ك)

$$= \text{مجـ } (\text{ص}-\bar{\text{ص}})^2 = \text{مجـ ص}^2 - (\text{مجـ ص})^2/\text{ن}$$

التغير المفسر بسبب وجود المتغير المستقل = مجموع المربعات بسبب الانحدار (م.م.ب)

$$= \text{مجـ } (\hat{\text{ص}} - \bar{\text{ص}})^2$$

$$= \text{ب مجـ } (\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})$$

$$= \text{ب } (\text{مجـ س ص} - \text{مجـ س} \times \text{مجـ ص}/\text{ن})$$

التغيير غير المفسر والذي يرجع إلى عوامل عشوائية = مجموع مربعات البواقى (م.م.د) $= \text{مجـ } (\text{ص} - \hat{\text{ص}})^2$

$$\therefore \text{م.م.ك} = \text{م.م.ب} + \text{م.م.د}$$

وباســترجاع أســلوب تحلــيل التبايــن من الباب السابق ، نجد أن كل مجموع مــربعات يرتبط دائما بدرجات حرية (د . ح) معينة ، فمثلا مجموع المربعات الكلــي (م.م.ك) يرتــبط بعدد (ن-١) من درجات الحرية ، (تم خصم درجة حرية واحدة من حجم العينة مقابل تقدير $\bar{\text{ص}}$) ، بينما مجموع مربعات البواقي (م . م . د) فلــه (ن - ٢) مــن درجات الحرية (تم خصم درجتي حرية من حجم العينة مقابل تقدير المعلمتين γ ، β) ، وحيث أن :

د . ح (م.م . ك) = د . ح (م .م. ب) + د. ح (م.م . د)

ن-١ = د . ح (م . م. ب) + (ن -٢)

∴ درجــات الحــرية لمجموع المربعات بسبب علاقة الانحدار تساوي واحد (وهــي فــي الواقــع تســاوي عدد المتغيرات المستقلة في نموذج الانحدار) .

وبترتيــب مجامــيع المربعات السابقة مقترنة بدرجات الحرية في جدول يسمى جدول تحليل التباين Analysis of Variance نصل إلى الصورة التالية .

جدول تحليل التباين

مصدر الاختلاف	مجموع المربعات (م.م)	درجات الحرية د . ح	متوسط مجموع المربعات (م.م.م)
راجعا إلى الانحدار (م.م.ب)	مجـ(ص^ − ص̄)^{٢}= ب مجـ (س−س̄) × (ص−ص̄)	١	مجـ(ص^−ص̄)^{٢} /١
راجعا إلى البواقي م.م.د	مجـ (ص − ص^)^{٢}	ن−٢	مجـ (ص−ص^)^{٢} / (ن − ٢) = خ$^{٢}_{ص}$
م . م . ك	مجـ (ص − ص̄)^{٢}	ن−١	

ولاختبار الفرض العدمي القائل بأن β = صفر ، أي لا توجد علاقة خطية بين ص ، س أي لا يوجـد تأثير للمتغير المستقل س على المتغير التابع ص ، فإننا نقـارن قيمة متوسط مجموع المربعات بسبب الانحدار مع قيمة متوسط مجموع المـربعات للبواقـي ، حيث يكون خارج القسمة لهما هو متغير عشوائي يتبع توزيع ف :

$$\text{ف} = \frac{\text{مجـ (ص^ − ص̄)}^{٢}}{\text{مجـ (ص−ص^)}^{٢} \text{ / (ن −٢)}} = \frac{\text{ب مجـ (س − س̄) (ص−ص̄)}}{\text{خ}^{٢}_{\text{ص}}} \quad ...(٥٢)$$

حيـث خ$^{٢}_{ص}$ هو تباين تقدير معادلة خط الانحدار أو تباين البواقي ، وبمقارنة القـيمة العددية للمتغير ف مع قيمة جدولية مستخرجة من جدول توزيع ف عند درجـات حـرية (١ ، ن − ٢) ومستوى معنوية (α / ٢) ، يمكن اتخاذ قرار بقـبول أو رفـض الفرض العدمي . بالطبع – وكما هو متوقع – تكون نتيجة القـرار هـنا هي نفس النتيجة التي نصل لها سواء استخدمنا اختبار معنوية β باستخدام توزيع (ت) أو باستخدام فترة الثقة للمعلمة β .

ملحوظة :

١- عند ثبات درجات الحرية ومستوى المعنوية ، فإن مربع قيمة ت الجدولية = قيمة ف الجدولية .

٢- مـــن الممكن أن نصل إلى صيغة أخرى لوسيلة الاختبار ف بدلالة معامل التحديد ر^{2} على النحو التالي :

$$\therefore \text{ر}^{2} = \frac{\text{مجـ}\,(\hat{\text{ص}} - \bar{\text{ص}})^{2}}{\text{مجـ}\,(\text{ص} - \bar{\text{ص}})^{2}} \quad \& \quad 1 - \text{ر}^{2} = \frac{\text{مجـ}\,(\text{ص} - \hat{\text{ص}})^{2}}{\text{مجـ}\,(\text{ص} - \bar{\text{ص}})^{2}}$$

$$\therefore \text{ف} = \frac{\text{مجـ}\,(\hat{\text{ص}} - \bar{\text{ص}})^{2} \times (\text{ن} - 2)}{\text{مجـ}\,(\text{ص} - \hat{\text{ص}})^{2}}$$

بقسمة كل من البسط والمقام على $\text{مجـ}\,(\text{ص} - \bar{\text{ص}})^{2}$ نصل إلى :

$$\boxed{\text{ف} = \frac{\text{ر}^{2}}{1 - \text{ر}^{2}} \times (\text{ن} - 2)} \qquad ...(٥٣)$$

أى أنـــه إذا علمـــت قـــيمة ر^{2} ، أمكن الوصول إلى قيمة وسيلة الاختبار (ف) والعكس صحيح .

مثال (١٥)

بفرض أنك حصلت على البيانات التالية :

س	١	٢	٣	٤	٥	٦
ص	٢	٤	٤	٦	٩	١٠

المطلوب :

١- معادلة خط انحدار ص / س

٢- اختبار معنوية معادلة خط الانحدار أي اختبار β = صفر مستخدما في ذلك أســـلوب تحليل التباين عند α = ٥ % حيث القيمة الجدولية : ف (١ ، ٤ ، ٥ %) = ٧,٧١

٣- معامل التحديد ومعامل الارتباط .

الحل :

س	ص	$س^2$	$ص^2$	س ص
١	٢	١	٤	٢
٢	٤	٤	١٦	٩
٣	٤	٩	١٦	١٢
٤	٦	١٦	٣٦	٢٤
٥	٩	٢٥	٨١	٤٥
٦	١٠	٣٦	١٠٠	٦٠
٢١	٣٥	٩١	٢٥٣	١٥١

١- معادلة خط انحدار ص / س : ص = أ + ب س ، حيث

ب = (مجـ س ص − مجـ س × مجـ ص / ن) / (مجـ س٢ − (مجـ س)٢ / ن) = (١٥١-٢١×٣٥ / ٦) / (٩١-(٢١)٢/٦)

= (١٥١-١٢٢,٥) / (٩١-٧٣,٥) = ٢٨,٥ / ١٧,٥ = ١,٦٢٨

أ = مجـ ص / ن − ب × مجـ س / ن = ٣٥ / ٦ − ١,٦٢٨ × ٢١ / ٦

= ٥,٨٣٣ − ٥,٦٩٨ = ٠,١٣٥

∴ ص = ٠,١٣٥ + ١,٦٢٨ س

٢- اختبار β = صفر عن طريق تحليل التباين

يقتضي جدول تحليل التباين حساب الكميات التالية :

م . م . ك = مجـ ص٢ − (مجـ ص)٢ / ن = ٢٥٣ − (٣٥)٢/٦

= ٢٥٣ – ٢٠٤,١٧ = ٤٨,٨٣

م . م . بسبب الانحدار = ب × (مجـ س ص – مجـ س × مجـ ص / ن)

= ١,٦٢٨ (١٥١ – ١٢٢,٥) = ١,٦٢٨ × ٢٨,٥ =٤٦,٣٩٨

جدول تحليل التباين

المصدر	م . م	د. ح	م . م . م	نسبة التباين ف
م . م بسبب انحدار ص / س	٤٦,٣٩٨	١	٤٦,٣٩٨	٧٦,٣١٢٥
م .م. د (البواقي)	٢,٤٣٢	٤	٠,٦٠٨	
م.م.ك	٤٨,٨٣٠	٥		

خطوات الاختبار :

١- الفرض العدمي β = صفر (لا توجد علاقة حقيقية بين س ، ص)

٢- الفرض البديل β = صفر (توجد علاقة حقيقية بين س ، ص)

٣- ف المحسوبة = ٧٦,٣١٢٥

٤- ف الجدولية : ف (١، ٤ ، ٥ %) = ٧,٧١

٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .

٦- القرار : رفض الفرض العدمي ومن ثم قبول الفرض البديل

∴ β ≠ صفر أي هناك علاقة معنوية بين ص ، س

٣- معامل التحديد ر^{2} = $\frac{\text{م.م بسبب الانحدار}}{\text{م . م . ك}}$ = $\frac{٤٦,٣٩٨}{٤٨,٨٣}$ = ٠,٩٥٠٢

أي أن المتغـير المسـتقل يؤثر في المتغير التابع بنسبة ٩٥% تقريبا ، والباقي يـرجع لتأثير العوامل العشوائية أو بمعنى آخر ٩٥% من الاختلافات الكلية في التابع ص ، تـرجع إلى وجود المتغير المستقل س ، أما معامل الارتباط فهو

الجذر التربيعي لمعامل التحديد واشارة معامل الارتباط هي نفس إشارة ميل خط الانحدار (ب) .

∴ معـــامل الارتــباط = $\sqrt{٠,٩٥٠٢}$ = ٠,٩٧٤٧ ، وحيث أن إشارة ميل خط الانحدار (ب) موجبة ، يكون هناك إرتباط قوي وموجب بين (ص ، س) .

(١١) استخدام نموذج الانحدار في عملية التنبؤ بفترة ثقة :

بعــد أن توصلنا إلى وصف العلاقة بين المتغير التابع (ص) والمستقل (س) والتــي على الصورة : ص^ = أ + ب س ، وبعــد أن تأكدنا من صحة العلاقة بينهما عن طريق اختبار β= صفر ، يتبقى لنا كيفية استخدام تلك العلاقة الانحدارية في التنبؤ بقيمة المتغير التابع (ص) عند قيمة محددة معلومة للمتغير المستقل (س) ، هنا ينشأ نوعين من التنبؤ :

(١) التنبؤ بمتوسط قيم (ص) عند قيمة محددة معلومة للمتغير س .

(٢) التنبؤ بقيمة (ص) عند قيمة محددة معلومة للمتغير س .

للتوضيح ، في مثال (٣) كانت العلاقة بين كمية الانتاج من القمح (ص) وكمية السـماد المســتخدم (س) على الصورة : ص = ٢٧,١٢ + ١,٦٦ س . هنا قد نرغــب فــي تقدير متوسط كمية الانتاج لعدد من الأفدنة عندما يستخدم في كل فــدان ١٥ كـيلو سماد مثلاً أي س = ١٥ أو قد نرغب في تقدير انتاج فدان واحد فقط عندما يعطي كمية سماد س = ١٥ وحدة .

لكن أي التقديرين : تقدير متوسط قيم $\hat{ص}$ لعدة أفدنة أو تقدير قيمة $\hat{ص}$ في فدان واحد ، يحقق دقة أكبر ؟ قبل الإجابة على ذلك ، دعنا أولا نتنبأ بتلك القيم .

(أ) تقدير متوسط الإنتاج ($\hat{ص}$) لعدة فترات عندما تكون س في كل فترة = ١٥ هي :

$\hat{ص}$ = ٢٧,١٢ + ١,٦٦ × ١٥ = ٥٢,٠٢ كيلو جرام

(ب) تقدير كمية الانتاج ($\hat{ص}$) عندما تكون س = ١٥ هي

$\hat{ص}$ = ٢٧,١٢ + ١,٦٦ × ١٥ = ٥٢,٠٢ كيلو جرام

من الواضح أن كلا التقديرين : متوسط ($\hat{ص}$) ، $\hat{ص}$ هما نفس القيمة عند س = ١٥ ، (كــلا التقديرين يعتبرا تقديرا بنقطة أو تقدير وحيد القيمة) ، لكن الفرق بيــنهما يكمن في الدقة النسبية لكل منهما ، وهذه الدقة النسبية تقاس عن طريق خطأ المعاينة أو الخطأ المعيارى لكل تقدير ، وهي على الصورة التالية:

أ- تقديــر الخطــأ المعياري للمقدر $\hat{ص}$ ، حيث $\hat{ص}$ هى متوسط قيم $\hat{ص}$ عند قيمة محددة معلومة للمتغير س ولتكن س. هو :

$$ع(\hat{ص}) = خ_ص \sqrt{\frac{١}{ن} + \frac{(س. - \bar{س})^٢}{مجـ (س - \bar{س})^٢}} \quad ... (٥٤)$$

حيث $خ_ص$ هو الخطأ المعياري لمعادلة خط الانحدار = $\frac{مجـ (ص - \hat{ص})^٢}{ن - ٢}$

ب- تقديــر الخطــأ المعياري للمقدر $\hat{ص}$، حيث $\hat{ص}$ هى قيمة $\hat{ص}$ عند قيمة محددة معلومة للمتغير س ولتكن س. هو :

$$ع(\hat{ص}) = خ_ص \sqrt{١ + \frac{١}{ن} + \frac{(س. - \bar{س})^٢}{مجـ (س - \bar{س})^٢}} \quad ... (٥٥)$$

وعن طريق الأخطاء المعيارية يمكن إنشاء فترة ثقة للتقديرات ($\hat{ص}$) سواء أكان التقدير هو متوسط قيم $\hat{ص}$ أو قيمة $\hat{ص}$ المفردة على النحو التالى :

$$\mu_{ص^\wedge / س.} = \hat{ص} \pm ت_{(ن-٢ ، \alpha/٢)} \times ع(\hat{ص}) \quad ...(٥٦)$$

حيــث : $\hat{ص}$ = أ + ب س. ، ع ($\hat{ص}$) إمــا أن تكون المعادلة (٥٤) فى حالة تقدير متوسط $\hat{ص}$ أو تكون المعادلة (٥٥) فى حالة تقدير قيمة وحيدة لــ $\hat{ص}$.

واستكمالاً للمثال المطروح بين أيدينا نعلم من هذا المثال أن :

ن = ١٠ ، $\bar{س}$ = ١٨٠÷١٠ = ١٨ ، خ(ص) = ٢,٤٣ ، مجـ(س - $\bar{س}$)^٢ = ٥٧٦

والمطلوب هو :

١- تقدير متوسط $\hat{\text{ص}}$ بفترة ثقة ٩٥% عند س. = ١٥ .

٢- تقدير $\hat{\text{ص}}$ بفترة ثقة ٩٥% عند س. = ١٥ .

بالنسبة للمطلوب الأول : تقدير متوسط $\hat{\text{ص}}$ بفترة ثقة ٩٥% :

$$\mu_{\text{ص}^\wedge/\text{س.}} = \hat{\text{ص}} \pm \text{ت}_{(\text{ن}-٢\ ،\ \alpha/٢)} \times \text{ع}\,(\hat{\text{ص}})$$

$$= (\text{أ} + \text{ب س}) \pm \text{ت}_{(١٠-٢\ .\ ٠,٠٢٥)} \times \text{خص} \sqrt{\frac{١}{\text{ن}} + \frac{(\text{س.}-\bar{\text{س}})^٢}{\text{مجـ س}^٢ - (\text{مجـ س})^٢ / \text{ن}}}$$

$$= (٢٧,١٢ + ١,٦٦٦ \times ١٥) \pm \text{ت}_{(٨\ ،\ ٠,٠٢٥)} \times ٢,٤٣ \sqrt{\frac{١}{١٠} + \frac{(١٥-١٨)^٢}{٥٧٦}}$$

$$= ٥٢,٠٢ \pm ٢,٣٠٦ \times ٠,٨٢٦ = ٥٢,٠٢ \pm ١,٩١ = ٥٣,٩٣ \ \&\ ٥٠,١١$$

أى أنــه عندما يتم استخدام كمية السماد س = ١٥ وحدة فى العديد من الأفدنة ، فإن متوسط إنتاج الفدان المتوقع سيتراوح بين ٥٠,١١ وحدة ، ٥٣,٩٣ وحدة بدرجة ثقة ٩٥% (لاحظ أن مدى الثقة هنا= ٥٣,٩٣ -٥٠,١١= ٣,٨٢) .

وبالنسبة للمطلوب الثانى : تقدير قيمة ص^ بفترة ثقة ٩٥% :

$$\mu_{\text{ص}^\wedge/\text{س.}} = \hat{\text{ص}} \pm \text{ت}_{(\text{ن}-٢\ ،\ \alpha/٢)} \times \text{ع}\,(\hat{\text{ص}})$$

$$= ٥٢,٠٢ \pm ٢,٣٠٦ \times \text{خص} \sqrt{١ + \frac{١}{\text{ن}} + \frac{(\text{س.}-\bar{\text{س}})^٢}{\text{مجـ}\,(\text{س}-\text{س})^٢}}$$

$$= ٥٢,٠٢ \pm ٢,٣٠٦ \times ٢,٤٣ \times \sqrt{١ + \frac{١}{١٠} + \frac{(١٥-١٨)^٢}{٥٧٦}}$$

$$= ٥٢,٠٢ \pm ٢,٣٠٦ \times ٢,٥٦٧ = ٥٢,٠٢ \pm ٥,٩٢ = ٥٧,٩٤ \ \&\ ٤٦,١٠$$

أى أنــه عــند استخدام كمية السماد س=١٥ وحدة فى فدان واحد ، فإن إنــتاجه المــتوقع سيتــراوح بيــن ٤٦,١ ، ٥٧,٩٤ وحــدة بدرجة ثقة ٩٥% (لاحظ أن مدى الثقة هنا = ٥٧,٩٤ - ٤٦,١٠ = ١١,٨٤ .

من مقارنة فترة الثقة لكلا التقديرين ، يمكن أن نتوصل للنتائج التالية :

١- بصـــفة عامة كلما زاد حجم العينة ، كلما قل الفرق بين حدى الثقة ، وكلما اقــترب الـــتقدير مـــن القيمة الحقيقية المجهولة وهذا يعنى زيادة الدقة فى التقدير.

٢- اتساع فترة الثقة عند تقدير صٌ^ عن فترة الثقة عند تقدير متوسط صٌ^ ، وهذا يعني أن الخطأ المعياري فى تقدير صٌ^ أكبر من الخطأ المعياري فى تقدير متوسط صٌ^ ، وهذا واضح إذ أن ع(صٌ^) فى الحالة الأولى = ٢,٥٦٧ بينما ع(صٌ^) فى حالة المتوسط =٠,٨٢٦ .

٣- الخطأ المعياري فى تقدير متوسط صٌ^ أو فى تقدير صٌ^ يصل إلى أدنى قيمة له عندما تكون س. = سَ .

٤- كلما ابتعدت س. عن سَ كلما زاد الخطأ المعياري فى تقدير صٌ^ أو متوسط صٌ^ والعكـــس صــــحيح ، لذا ينصح أن يتم التنبؤ بقيمة صٌ^ عند قيمة س. القريبة بقدر الإمكان من سَ .

٥- إذا ابـــتعدت س. عن سَ بدرجة كافية ووقعت خارج مجال قيم س بالعينة ، يكون من الخطر عمل أى استنتاجات حول صٌ^ أو متوسط صٌ^ .

مثال (١٦) (مثال شامل):

البيانات التالية تبين قيمة الإنفاق على الدعاية بالألف جنيه (س) وقيمة المبيعات بالألف جنيه (ص) لإحدى الشركات خلال خمسة شهور متتالية:

الشهر	يناير	فبراير	مارس	إبريل	مايو
الإنفاق (س)	١	٢	٣	٤	٥
المبيعات (ص)	١	١	٢	٢	٤

<u>المطلوب :</u>

١- ارسم الشكل الانتشارى لهذه البيانات ثم وفق لها أفضل خط مستقيم مستخدماً طريقة المربعات الصغرى .

٢- أوجد قيمة معامل التحديد ومعامل الارتباط .

٣- قدر الخطأ المعياري لنموذج خط الانحدار .

٤- قدر بفترة ثقة ٩٥% ميل خط الانحدار β .

٥- اختبار صلاحية نموذج خط الانحدار فى الحالات التالية :

(أ) β = صفر (ب) β = ٠,٣ حيث α = ٥% .

مستخدماً فى الحالتين أسلوب فترة الثقة ثم أسلوب اختبار المعنوية .

٦- اختبار صلاحية نموذج خط الانحدار ، مستخدماً فى ذلك أسلوب تحليل التباين عند مستوى معنوية ٥% .

٧- قدر بفترة ثقة ٩٥% ما يلي :

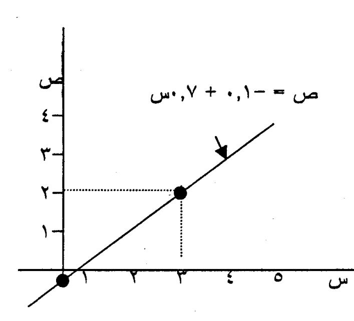
أ- متوسط قيمة المبيعات فى الشهور القادمة عندما تكون قيمة الدعاية فى أى شهر س. = ٤ (ألف جنيه) .

ب- قـيمة المبيعات فى الشهر القادم عندما تكون قيمة الدعاية فى ذلك الشهر س.=٤ (ألف جنيه) .

ج- أعـد المطلوب (أ) مرة أخرى عند س. = ٣ ثم عند س. = ٧ ثم علق على النتائج.

الحل:

س	ص	س٢	ص٢	س ص
١	١	١	١	١
٢	١	٤	١	٢
٣	٢	٩	٤	٦
٤	٢	١٦	٤	٨
٥	٤	٢٥	١٦	٢٠
١٥	١٠	٥٥	٢٦	٣٧



<u>١ - الشكل الانتشارى وأفضل خط مستقيم .</u>

عـــند رصد البيانات بيانياً نلاحظ أنه يمكن استبدالها تقريباً بخط مستقيم يــتجه صــعوداً من أسفل اليسار إلى أعلى اليمين ، ولرسم أفضل خط مستقيم بطريقة المربعات الصغرى نجد أن :

ص^ = أ + ب س حيث

$$\text{ب} = \frac{\text{مجـ س ص} - \text{مجـ س} \times \text{مجـ ص} / \text{ن}}{\text{مجـ س}^{٢} - (\text{مجـ س})^{٢} / \text{ن}} = \frac{٣٧ - ١٥ \times ١٠/٥}{٥٥ - (١٥)^{٢}/٥} = \frac{٧}{١٠} = ٠,٧$$

(وتسمي ب بميل خط الانحدار أو معدل تغير ص مع س عندما تتغير س بوحدة واحدة).

$$\text{أ} = \frac{\text{مجـ ص}}{\text{ن}} - \text{ب} \times \frac{\text{مجـ س}}{\text{ن}} = \frac{١٠}{٥} - ٠,٧ \times \frac{١٥}{٥} = ٢ - ٢,١ = -٠,١$$

(وتسمي أ بالجزء المقطوع من المحور الرأسي وهى مقدار ثابت ، أى قيمة ص عندما تكون س=صفر) .

∴ معادلة خط الانحدار تصبح على الصورة : **ص^ = -٠,١ + ٠,٧ س**

ولرسم هذا الخط نكتفى برصد نقطتين بيانياً فمثلاً :

عند س = صفر فإن ص = -٠,١ ∴ أول نقطة (٠ ، -٠,١)

عند س = ٣ فإن ص = ٢ ∴ ثانى نقطة (٣،٢)

∴ عن طريق رصد نقطتين (٠،-٠,١) ، (٣،٢) والتوصيل بينهما بخط مستقيم ، نحصل على أفضل خط انحدار يمثل تلك البيانات .

<u>٢ - معامل التحديد ومعامل الارتباط :</u>

$$\text{معامل التحديد ر}^{٢} = \frac{\text{التغير المفسر}}{\text{التغير الكلى}} = \frac{\text{مجـ }(\text{ص}^{\wedge} - \bar{\text{ص}})^{٢}}{\text{مجـ }(\text{ص} - \text{ص})^{٢}}$$

= ب (مجـ س ص – مجـ س × مجـ ص / ن) ÷ (مجـ ص^{٢} – (مجـ ص)^{٢} / ن)

= ٠,٧ (٣٧ – ١٥ × ٥/١٠) ÷ (٢٦ – (١٠)^{٢} /٥) = (٠,٧ × ٧) ÷ ٦ = ٠,٨١٦٦

أى أن ٨١,٦٦% مـــن التغير الكلى فى المبيعات (ص) ترجع إلى الإنفاق على الدعاية ، وأن الباقى ١٨,٣٤% ترجع إلى عوامل عشوائية ، معامل الارتباط هو الجذر التربيعي لمعامل التحديد أما إشارته (موجب أو سالب) فتحدد وفق إشارة ميل خط الانحدار ب .

ر = √ ر^{٢} = √٠,٨١٦٦ = ٠,٩ .

وحيــث أن إشــارة (ب) هـــى موجبة ، تكون إشارة معامل الارتباط هى ايضاً موجــبة ، أى أن هــناك ارتــباط طردى (موجب) قوى بين قيمة الإنفاق على الدعاية وقيمة المبيعات المتحققة .

٣- الخطأ المعياري لنموذج خط الانحدار :

خ$_{ص}$ = √ [مجـ (ص – ص^)^{٢} ÷ (ن – ٢)] أو خ$_{ص}$ = √ [(مجـ ص^{٢} – أ مجـ ص – ب مجـ س ص) ÷ (ن – ٢)]

ويمكن حساب قيمة خ$_{ص}$ بأى صيغة من تلك الصيغ .

س	ص	ص^=–٠,١+٠,٧س	ص–ص^	(ص–ص^)^{٢}
١	١	٠,٦	١–٠,٦= ٠,٤	٠,١٦
٢	١	١,٣	١–١,٣= –٠,٣	٠,٠٩
٣	٢	٢	٢–٢ = صفر	صفر
٤	٢	٢,٧	٢–٢,٧= –٠,٧	٠,٤٩
٥	٤	٣,٤	٤–٣,٤= ٠,٦	٠,٣٦
			صفر	١,١٠

$$\text{خ}_{\text{ص}} = \sqrt{\frac{\text{مجـ}(\text{ص} - \hat{\text{ص}})^{٢}}{\text{ن} - ٢}} = \sqrt{\frac{١,١}{٣}} = \sqrt{٠,٣٦٦} = ٠,٦١$$

$$\text{أو } \text{خ}_{\text{ص}} = \sqrt{\frac{\text{مجـ ص}^{٢} - \text{أ مجـ ص} - \text{ب مجـ س ص}}{\text{ن} - ٢}} = \sqrt{\frac{٢٦ - (-٠,١) \times ١٠ - ٠,٧ (٣٧)}{٥ - ٢}}$$

$$= \sqrt{\frac{٢٦ + ١ - ٢٥,٩}{٣}} = \sqrt{\frac{١,١}{٣}} = \sqrt{٠,٣٦٦} = ٠,٦١$$

<u>٤- تقدير ميل خط الانحدار β بفترة ثقة ٩٥% :</u>

$\beta = \text{ب} \pm \text{ت}_{(\text{ن}-٢ \,،\, \alpha/٢)} \times \text{ع}_{(\text{ب})}$ ، (ع(ب) : الخطأ المعياري للتقدير ب)

$$= \text{ب} \pm \text{ت}_{(\text{ن}-٢ \,،\, \alpha/٢)} \times \frac{\text{خ}_{\text{ص}}}{\sqrt{\text{مجـ س}^{٢} - (\text{مجـ س})^{٢} / \text{ن}}}$$

$$= ٠,٧ \pm \text{ت}_{(٣ \,،\, ٠,٠٢٥)} \times \frac{٠,٦١}{\sqrt{٥٥ - (١٥)^{٢} / ٥}}$$

$$= ٠,٧ \pm ٣,١٨٢ \times ٠,١٩ = ٠,٧ \pm ٠,٦١ = ١,٣١ \text{ ، } ٠,٠٩$$

∴ معامل خط الانحدار β تتراوح قيمته بين ٠,٠٩ ، ١,٣١ بدرجة ثقة ٩٥% .

<u>٥- اختبار صلاحية نموذج خط الانحدار :</u>

أولاً : باستخدام أسلوب فترة الثقة :

$$\therefore \beta = \text{ب} \pm \text{ت}_{(\text{ن}-٢ \,،\, \alpha/٢)} \times \text{ع(ب)} = ٠,٧ \pm ٠,٦١ = ١,٣١ \text{ ، } ٠,٠٩$$

أ- عندما يكون الفرض العدمى هو β = صفر

حيث أن الصفر يقع خارج حدى الثقة ، فهذا يعني رفض الفرض العدمى بأن β = صفر ومن ثم قبول الفرض البديل أى β ≠ صفر وهذا يعني وجود تأثير معنوي للمتغير المستقل على المتغير التابع .

ب- عندما يكون الفرض العدمى هو β = ٠,٣

حيـــث أن القيمة المراد اختبارها للفرض العدمى وهو β = ٠,٣ تــقع داخل حـــدى الثــقة ، فهـــذا يعنـــي قبول نص الفرض العدمى أى أن β = ٠,٣ .

<u>ثانياً : باستخدام أسلوب اختبار المعنوية :</u>

أ- عندما يكون الفرض العدمى هو β = صفر

<u>خطوات الاختبار :</u>

١- الفرض العدمى : β = صفر

٢- الفرض البديل : β ≠ صفر

٣- $\text{ت} = \frac{\text{ب} - \beta}{\text{ع}_{(\text{ب})}} = \frac{\text{٠,٧} - \text{صفر}}{\text{٠,١٩}} = \text{٣,٦٨٠}$

٤- ت الجدولية : $\text{ت}_{(\text{ن}-٢ \text{ ، } \alpha/٢)} = \text{ت}_{(٣ \text{ ، } ٠,٠٢٥)} = ٣,١٨٢$

٥- المقارنة: ت المحسوبة أكبر من ت الجدولية .

٦- القـــرار: رفض الفرض العدمى وقبول الفرض البديل.

∴ β ≠ صفر وهى نفس النتيجة التى توصلنا إليها عن طريق فترة الثقة.

ب- عندما يكون الفرض العدمى هو β = ٠,٣

<u>خطوات الاختبار :</u>

١- الفرض العدمى β = ٠,٣

٢- الفرض البديل β ≠ ٠,٣

٣- $\text{ت} = \frac{\text{ب} - \beta}{\text{ع}_{(\text{ب})}} = \frac{٠,٧ - ٠,٣}{٠,١٩} = ٢,١١$

٤- ت الجدولية : $\text{ت}_{(\text{ن}-٢ \text{ ، } \alpha/٢)} = \text{ت}_{(٣ \text{ ، } ٠,٠٢٥)} = ٣,١٨٢$

٥- المقارنة : ت المحسوبة اقل من ت الجدولية .

٦- القــرار : قبول الفرض العدمى ، وهى نفس النتيجة التى توصلنا إليها عن طريق فترات الثقة.

<u>٦- استخدام تحليل التباين فى تحليل الانحدار :</u>

لتكوين جدول تحليل التباين ، يلزم حساب المجاميع التالية :

م.م.ك = مجـ ص^{٢} −(مجـ ص)^{٢}/ ن = ٢٦−(١٠)^{٢}/٥ = ٦

م.م. بسبب الانحدار = ب × (مجـ س ص − مجـ س× مجـ ص/ن)

= ٠,٧ (٣٧ − ١٥×١٠/٥) = ٠,٧×٧ = ٤,٩

جدول تحليل التباين

المصدر	م.م	د . ح	م.م.م	ف
م.م. بسبب انحدار ص/س	٤,٩	١	٤,٩	١٣,٣٨٨
م.م.د . البواقى	١,١	٣	٠,٣٦٦	
م.م.ك	٦	٤		

<u>خطوات الاختبار :</u>

١- الفرض العدمى β = صفر

٢- الفرض البديل $\beta \neq$ صفر

٣- ف = ١٣,٣٨٨

٤- ف الجدولية = ف (١ ، ٣ ، ٥%) = ١٠,١٣

٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .

٦- القرار : رفض الغرض العدمى وقبول الفرض البديل .

∴ $\beta \neq$ صفر أى أن هناك تاثير للمتغير المستقل على المتغير التابع .

ملاحظات :

١- القـــرار الـــذى نتوصل إليه باستخدام تحليل التباين ، هو نفس القرار الذى نتوصل إليه سواء باستخدام فترة الثقة أو باستخدام اختبار المعنوية للمعامل β عن طريق توزيع ت .

٢- بمعلومـــية قيمة ت الجدولية نصل إلى قيمة ف الجدولية او العكس ، حيث مـــربع قـــيمة ت الجدولـــية = قيمة ف الجدولية ، فمثلاً إذا كانت قيمة ف الجدولية = ١٠,١٣ فإن قيمة ت الجدولية = $\sqrt{١٠,١٣}$ = ٣,١٨٢ وهى نفس القيمة التى استخدمناها من قبل ، هذا على فرض ثبات حجم العينة ومستوى المعنوية .

٣- من جدول تحليل التباين يمكن استنتاج معامل التحديد :

$$ر^٢ = \frac{\text{م.م بسبب الانحدار}}{\text{م.م.ك}} = \frac{٤,٩}{٦} = ٠,٨١٦٦٦$$

ومعامل الارتباط = $\sqrt{٠,٨١٦٦}$ = ٠,٩

٤- من الممكن حساب قيمة ف بمعلومية معامل التحديد وفق العلاقة التالية :

$$ف = \frac{ر^٢}{١-ر^٢} \times \frac{ن-ك}{ك-١}$$ حيث ك: عدد المعالم المقدرة وهى هنا :

γ ، β أى أن ك =٢

$$ف = \frac{٠,٨١٦٦}{٠,١٨٣٤} \times \frac{٥-٢}{٢-١} = ١٣,٣٥٨$$

وهـــى نفس القيمة التى ظهرت فى جدول تحليل التباين (هناك اختلاف ضئيل بسبب تقريب العمليات الحسابية) .

٧- التنبؤ بفترة ثقة ٩٥% :

أ- تقدير متوسط المبيعات عند س. = ٤ بفترة ثقة ٩٥% .

$$\mu_{\hat{\text{ص}}/\text{س}_0} = \hat{\text{ص}} \pm \text{ت}_{(\text{ن}-٢\ ،\ \alpha/٢)} \times \text{ع}(\hat{\text{ص}})$$

$$= (\text{أ} + \text{ب س}_0) \pm \text{ت}_{(٣\ ،\ ٠,٠٢٥)} \times \text{خ}_{\text{ص}} \sqrt{\frac{١}{٥} + \frac{(\text{س}_0 - \bar{\text{س}})^٢}{\text{مجـ س}^٢ - (\text{مجـ س})^٢/\text{ن}}}$$

من هذا المثال توصلنا إلى :

$\text{أ} = -٠,١$ ، $\text{ب} = ٠,٧$ ، $\text{خ}_{\text{ص}} = ٠,٦١$ ، $\bar{\text{س}} = ٣$

$$\text{مجـ س}^٢ - (\text{مجـ س})^٢ / \text{ن} = ٥٥ - (١٥)^٢ / ٥ = ١٠$$

$$\mu_{\hat{\text{ص}}/\text{س}_0} = (-٠,١ + ٠,٧ \times ٤) \pm ٣,١٨٢ \times ٠,٦١ \times \sqrt{\frac{١}{٥} + \frac{(٤-٣)^٢}{١٠}}$$

$$= ٢,٧٠ \pm ٣,١٨٢ \times ٠,٣٣ = ٢,٧ \pm ١,٠٦ = ٣,٧٦ \ ،\ ١,٦٤$$

أى أنه عندما تقوم الشركة بإنفاق ٤ آلاف جنيه على الدعاية شهرياً فإن متوسط قيمة المبيعات المتوقعة ستتراوح بين ١,٦٤ ، ٣,٧٦ ألف جنيه .

ب- تقدير قيمة المبيعات ص^ عند س. = ٤ بفترة ثقة ٩٥% :

$$\mu_{\hat{\text{ص}}/\text{س}_0} = \hat{\text{ص}} \pm \text{ت}_{(\text{ن}-٢\ ،\ \alpha/٢)} \times \text{ع}(\hat{\text{ص}})$$

$$= ٢,٧ \pm ٣,١٨٢ \times \text{خ}_{\text{ص}} \times \sqrt{١ + \frac{١}{\text{ن}} + \frac{(\text{س}_0 - \bar{\text{س}})^٢}{\text{مجـ س}^٢ - (\text{مجـ س})^٢/\text{ن}}}$$

$$= ٢,٧ \pm ٣,١٨٢ \times ٠,٦١ \times \sqrt{١ + \frac{١}{٥} + \frac{(٤-٣)^٢}{١٠}}$$

$$= ٢,٧ \pm ٣,١٨٢ \times ٠,٧$$

$$= ٢,٧ \pm ٢,٢٢ = ٤,٩٢ \ ،\ ٠,٤٨$$

أى أنه عندما تنفق الشركة ٤ آلاف جنيه فى الشهر القادم فإن المبيعات المتوقعة فى ذلك الشهر ستتراوح بين ٠,٤٨ ، ٤,٩٢ ألف جنيه .

ج- تقدير متوسط المبيعات عند س. = ٣ ثم عند س. = ٧

$$\mu_{\hat{\text{ص}}/\text{س}_0} = (-٠,١ + ٠,٧\text{س}_0) \pm \text{ت}_{(٣\ ،\ ٠,٠٢٥)} \times \text{ع}(\hat{\text{ص}})$$

$$= (-٠,١ + ٠,٧ \times ٣) \pm ٣,١٨٢ \times \text{خ}_{\text{ص}} \sqrt{\frac{١}{\text{ن}} + \frac{(\text{س}_0 - \bar{\text{س}})^٢}{\text{مجـ}\ (\text{س} - \bar{\text{س}})^٢}}$$

$$= ٢ \pm ٣,١٨٢ \times ٠,٦١ \times \sqrt{\frac{١}{٥} + \frac{(٣-٣)^{٢}}{١٠}}$$

$$= ٢ \pm ٣,١٨٢ \times ٠,٢٧ = ٢ \pm ٠,٨٦ = ٢,٨٦ \ \& \ ١,١٤$$

وعند س. = ٧ ، نجد أن :

$$\mu_{ص^\wedge/س.} = ٧ = ٤,٨ \pm ٣,١٨٢ \times ٠,١٦ \times \sqrt{\frac{١}{٥} + \frac{(٧-٣)^{٢}}{١٠}}$$

$$= ٤,٨ \pm ٣,١٨٢ \times ٠,٨٢ = ٤,٨ \pm ٢,٦ = (٢,٢ ، ٧,٤)$$

ملحوظة :

إذا كان الحد الأدنى لفترة الثقة سالب القيمة ، فإننا نعتبره صفراً ، لأنه من غير المنطقى أن تكون قيمة المبيعات سالبة القيمة .

تعليق :

يلاحظ أنه عندما كانت س. = $\bar{س}$ = ٣ كان الخطأ المعياري ع(ص^) = ٠,٢٧ وعندما كانت س. = ٤ أصبح الخطأ المعياري ٠,٣٣ وعندما كانت س. = ٧ ارتفع الخطأ المعياري إلى ٠,٨٢ أى أنه كلما ابتعدت س. عن $\bar{س}$ كلما زاد الخطـأ المعـياري ، ويزداد الخطأ المعياري بصورة خطيرة إذا كانت س. تقع خارج مجال قيم س فى العينة (لاحظ أن مجال س فى العينة يتراوح بين ١ ،٥) لذا ينصح أن يتم التنبؤ بقيمة ص^ عند قيمة س. القريبة من متوسطها $\bar{س}$.

مثال (١٧) شامل :

قــام مدير إحدى الشركات بتحليل العلاقة بين مستوى أداء العاملين بها (ص) ومعدلهم التراكمى فى البكالوريوس (س) حيث يعتقد بوجود علاقة بينهم ، فحصل على النتائج التالية :

مسلسل	س	ص	س ص	س^2	ص^2
١	٣	٥	١٥	٩	٢٥
٢	٢	٤	٨	٤	١٦
٣	٤	٤	١٦	١٦	١٦
٤	١٢	٩	١٠٨	١٤٤	٨١
٥	١١	٨	٨٨	١٢١	٦٤
٦	٨	٩	٧٢	٦٤	٨١
٧	٩	٧	٦٣	٨١	٤٩
٨	٧	٨	٥٦	٤٩	٦٤
٩	٦	٥	٣٠	٣٦	٢٥
١٠	٥	٦	٣٠	٢٥	٣٦
١١	٤	٨	٣٢	١٦	٦٤
١٢	٨	٤	٣٢	٦٤	١٦
١٣	٣	٧	٢١	٩	٤٩
١٤	١٢	٦	٧٢	١٤٤	٣٦
١٥	٩	٨	٧٢	٨١	٦٤
١٦	٨	٥	٤٠	٦٤	٢٥
١٧	١١	١٠	١١٠	١٢١	١٠٠
١٨	٧	٧	٤٩	٤٩	٤٩
١٩	٨	٦	٤٨	٦٤	٣٦
٢٠	١٠	٥	٥٠	١٠٠	٢٥
ن=٢٠	مجـ س = ١٤٧	مجـ ص = ١٣١	مجـ س ص = ١٠١٢	مجـ س^2 = ١٢٦١	مجـ ص^2 = ٩٢١

(١) معادلة خط الانحدار ص = أ + ب س ، حيث :

مجـ ص = ن أ + ب مجـ س   ١٣١ = ٢٠أ + ١٤٧ ب

مجـ س ص = أ مجـ س + ب مجـ س^2 ١٠١٢ = ١٤٧أ + ١٢٦١ ب

وبحل تلك المعادلتين نجد أن : أ = ٤,٥٥ ، ب = ٠,٢٧

ويمكن استخدام صيغة القانون لحساب أ ، ب كما يلي :

$$ب = \frac{مجـ س ص - مجـ س × مجـ ص/ن}{مجـ س^{٢} - (مجـ س)^{٢}/ن} = \frac{١٠١٢ - ١٤٧×١٣١/٢٠}{١٢٦١ - (١٤٧)^{٢}/٢٠} = ٠,٢٧$$

$$أ = \bar{ص} - ب \bar{س} = \frac{١٣١}{٢٠} - ٠,٢٧ × \frac{١٤٧}{٢٠} = ٦,٥٥ - ١,٩٩ = ٤,٥٥$$

∴ معادلة خط الانحدار ص = ٤,٥٥ + ٠,٢٧ س

(٢) الخطأ المعياري للتقدير

$$خ_{ص} = \sqrt{(مجـ ص^{٢} - أ مجـ ص - ب مجـ س ص) ÷ (ن-٢)}$$

$$= \sqrt{(٩٢١ - ٤,٥٥ × ١٣١ - ٠,٢٧ × ١٠١٢) ÷ ١٨} = \sqrt{\frac{٥١,٧١}{١٨}}$$

$$= \sqrt{٢,٨٧} = ١,٦٩٥$$

(٣) فترة الثقة لمعامل الانحدار β عند ∝ = ٥% :

$$β = ب \pm ت_{(ن-٢ ، ∝/٢)} × ع(ب)$$

حيث ع (ب) : الخطأ المعياري للتقدير ب = $خ_{ص} ÷ \sqrt{مجـ س^{٢} - (مجـ س)^{٢}/ن}$

$$ع (ب) = \frac{١,٦٩٥}{\sqrt{١٢٦١ - (١٤٧)^{٢}/٢٠}} = \frac{١,٦٩٥}{١٣,٤٣٦٩} = ٠,١٢٦$$

$$β = ٠,٢٧ \pm ت_{(١٨ ، ٠,٠٢٥)} × ٠,١٢٦$$

$$= ٠,٢٧ \pm ٢,١٠١ × ٠,١٢٦ = ٠,٢٧ \pm ٠,٢٦٤٧ = ٠,٠٠٥٣ ، ٠,٥٣٤٧$$

(٤) اختبار β = صفر عند ∝ = ٥% باستخدام توزيع ت :

<u>**خطوات الاختبار :**</u>

١- الفرض العدمي β = صفر

٢- الفرض البديل β ≠ صفر

٣- $ت = \frac{ب-β}{ع(ب)} = \frac{ب}{ع(ب)} = \frac{٠,٢٧}{٠,١٢٦} = ٢,١٤٣$

٤- ت الجدولية : ت (١٨ ، ٠,٠٢٥) = ٢,١٠١

٥- المقارنة : ت المحسوبة أكبر من ت الجدولية .

٦- القرار : رفض الفرض العدمى وقبول الفرض البديل .

∴ β ≠ صفر،أى أن معامل انحدار المعدل التراكمى يختلف عن الصفر ويمكـــن استخدام فترة الثقة السابقة كأسلوب آخر لاختبار β = صفر ، وذلـــك بالبحث عن الصفر داخل حدى الثقة ، وحيث أن الصفر لا يقع داخـــل حـــدى الـــثقة يتخذ قرار يرفض الفرض العدمى وقبول الفرض البديل .

(٥) اختبار β = صفر باستخدام تحليل التباين :

م.م.ك = مجـــ ص٢ – (مجـــ ص)٢/ن = ٩٢١ –(١٣١)٢/٢٠ = ٦٣

م.م. بسبب انحدار ص/س = ب(مجـــ س ص –مجـــ س× مجـــ ص/ن)

= ٠,٢٧(١٠١٢ –١٤٧ ×١٣١/٢٠) =١٣,٢٧

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط مجموع المربعات	نسبة التباين	ف الجدولية
م.م. انحدار ص/س	١٣,٢٧	١	١٣,٢٧	٤,٨	ف(١، ١٨، ٥%) =٣,٠١
م.م.د (البواقى)	٤٩,٧٣	١٨	٢,٧٦٣		
م.م.ك	٦٣	١٩			

وحيث ان ف المحسوبة تتجاوز ف الجدولية عند مستوى المعنوية ٥% ، فهـــذا يعني رفض الفرض العدمي وقبول الفرض البديل ، أى أن : β ≠ صفر وهو نفس القرار الذي توصلنا إليه سابقاً .

(٦) تقدير متوسط أداء العاملين بفترة ثقة ٩٥% عند :

(أ) س = ٤ ، (ب) س = ٧,٣٥ ، (جـ) س = ١٠

بوضع كل قيمة من قيم س السابقة في المعادلة ص= ٤,٥٥+ ٠,٢٧ س نحصـــل علـــى قيمة ص^ المناظرة . ، ويبين الجدول التالي هذه القيم بالإضافة إلى قيم مقادير أخرى لازمة لحل هذا المطلوب ، مع العلم بأن

$\bar{س}$ = ١٤٧ ÷ ٢٠ = ٧,٣٥

س	ص^	ت $_{(١٨،٠,٠٢٥)}$	خ ص	$\sqrt{\frac{١}{ن} + \frac{(س_. - \bar{س})^٢}{مجـ س^٢ - (مجـ س)^٢ / ن}}$
٤	٥,٦٣	٢,١٠١	١,٦٩٥	$\sqrt{\frac{١}{٢٠} + \frac{١١،٢٢}{١٨٠,٥٥}}$ = ٠,٣٣٤٩
٧,٣٥	٦,٥٣	٢,١٠١	١,٦٩٥	$\sqrt{\frac{١}{٢٠} + \frac{صفر}{١٨٠,٥٥}}$ = ٠,٢٢٣٦
١٠	٧,٢٥	٢,١٠١	١,٦٩٥	$\sqrt{\frac{١}{٢٠} + \frac{٧,٠٢}{١٨٠,٥٥}}$ = ٠,٢٩٨٢

وبوضع هذه الكميات في العلاقة التالية :

$$\mu_{ص^/س.} = ص^ \pm ت_{(ن-٢ ، \alpha/٢)} \times ع(ص^)$$

نصل إلى فترات الثقة الثلاث التالية :

١- $\mu_{ص^/س.}$ = ٤,٤٣٧ & ٦,٨٢٣ مدى الثقة = ٢,٣٨٦

٢- $\mu_{ص^/س.}$ = ٥,٧٣٨ & ٧,٣٣١ مدى الثقة = ١,٥٩٣

٣- $\mu_{ص^/س.}$ = ٦,١٨٨ & ٨,٣١٢ مدى الثقة = ٢,١٢٤

يلاحظ أن فترة الثقة تزداد اتساعاً كلما ابتعدت س. عن $\bar{س}$ ، وهذا يعني أن تقدير المتوسط يقل الاعتماد كلما ابتعدت س. عن $\bar{س}$ ويرجع السبب في ذلك

إلـــى أن قـــيمة الخطأ المعياري للتقدير ع(ص^) تصل إلى أدنى قيمة لها عندما تصبح س. = س̄ وأن قيمته تزداد كلما ابتعدت س. عن س̄

(٧) تقدير قيمة ص^ (وليس متوسط ص^) بفترة ثقة ٩٥% :

عند س = ١٠ كانت ص^ = ٧,٢٥

$\mu_{\text{ص^/س.}}$ = ص^ ± ت $_{(\text{١٨ ، ٠,٠٢٥})}$ × ع (ص^)

$$= \text{٧,٢٥} \pm \text{٢,١٠١} \times \text{خ}_{\text{ص}} \sqrt{\text{١} + \frac{\text{١}}{\text{٢٠}} + \frac{(\text{١٠}-\text{٧,٣٥})^{\text{٢}}}{\text{١٨٠,٢٥}}}$$

= ٧,٢٥ ± ٢,١٠١ × ١,٦٩٥ × ١,٠٤٤

= ٧,٢٥ ± ٣,٧١٦ = ١٠,٩٦٦ ، ٣,٥٣٤ مدى الثقة =٧,٤٣٢

يلاحـــظ أن فترة الثقة الأخيرة أكثر إتساعا من أي فترة حصلنا عليها عند تقدير متوسط ص^

مثال (١٨) نصف محلول

بفرض أنك حصلت على البيانات التالية :

ص	٢٠	٢٨	٤٠	٤٥	٣٧	٥٢	٥٤	٤٣	٦٥	٥٦
س	٢	٣	٥	٤	٣	٥	٧	٦	٧	٨

المطلوب :

(أ) معادلة خط انحدار ص / س

الإجابة : ص = ١٤,٢٨ + ٥,٩٤ س

(ب) تباين تقديرات معادلة خط الانحدار والخطأ المعياري لها .

الاجابة : $\text{خ}^{\text{٢}}_{\text{ص}}$ = ٤٦,٩٧ ، $\text{خ}_{\text{ص}}$ = ٦,٨٥

(جـ) تباين التقدير (ب) والخطأ المعياري له .

الاجابة : $\text{ع}^{\text{٢}}$(ب) = ١,٣١ ، ع (ب) = ١,١٤

(حـ) اختبار β = صفر عند α = ٠,٠٥

الاجابة : β معنوية عند α = ٠,٠٥

(د) فترة الثقة للمعلمة β عند α = ٠,٠٥

الاجابة : β = ٣,٣١ ، ٨,٥٧

(هـ) معامل التحديد ومعامل الارتباط

الاجابة : ر^٢ = ٠,٧٧ ، ر = ٠,٨٨

مثال (١٩) (نصف محلول)

لتقدير أثر برنامج تدريبي على القدرة البيعية لمندوبي المبيعات باحدى الشركات ، جمعت البيانات التالية وهي تمثل (س) : الدرجة التي حصل عليها مندوب المبيعات بعد انتهاء البرنامج التدريبي الذي استمر شهرين ، (ص) : قيمة المبيعات لهم بالألف جنيه خلال السنة التالية لإنتهاء البرنامج التدريبي .

س	١٨	٢٦	٢٨	٣٤	٣٦	٤٢	٤٨	٥٢	٥٤	٦٠
ص	٥٤	٦٤	٥٤	٦٢	٦٨	٧٠	٧٦	٦٦	٧٦	٧٤

<u>المطلوب :</u>

(أ) : معادلة خط انحدار ص /س بطريقة المربعات الصغرى

ص = ٤٦,٥٢٤ + ٠,٤٩٩ س

(ب) قياس الخطأ المعياري لتقدير معادلة خط الانحدار مستخدما في ذلك كل من القيم الفعلية (ص) والقيم التقديرية (ص)

$خ^٢_{ص}$ = مجـ (ص−ص^)^٢ ÷ (ن −٢) = ١٧٠,٥٠١ ÷ ٨ = ٢١,٣٢

∴ $خ_{ص}$ = ٤,٦١٦

(جـ) اختبار β = صفر باستخدام تحليل التباين

م.م.ك = ٥٩٠,٤ ، م.م.انحدار ص/س = ٤١٩,٧٩٦

نسبة التباين (ف المسحوبة) = ١٩,٦٨٥ (تقريبا)، ف الجدولية

ف (١،٨،٥%)=٥,٣٢

القرار : رفض الفرض العدمي وقبول البديل أي $\beta \neq$ صفر

(د) معامل التحديد : ر٢ = ٠,٧١١ = ٧١,١% أما معامل الارتباط = ٠,٨٤٣

(هـ) اختبار β = صفر باستخدام توزيع ت

$$\text{ت المحسوبة} = \frac{\text{ب} - \beta}{\text{ع (ب)}} = \frac{\text{٠,٤٩٩ - صفر}}{\sqrt{\frac{\text{٢١,٣٢}}{\text{١٦٨٣,٦}}}} = \text{٤,٤١٦}$$

ت (٨ ، ٠,٠٢٥) = ٢,٣٠٦

القرار : رفض الفرض العدمي وقبول البديل

(و) تقدير β بفترة ثقة ٩٥%

β = ب $\pm$ ت $_{(\text{ن-٢ ، } \alpha/٢)}$ × ع (ب)

= ٠,٤٩٩ $\pm$ ٢,٣٠٦ × ٠,١٣٣ = ٠,٢٣٨ & ٠,٧٦

(ز) تقدير متوسط $\hat{\text{ص}}$ عند س. = ٣٩,٨ = $\bar{\text{س}}$

$$\text{ع}^2\ (\text{ص}^\wedge) = \text{٢١,٣٢} \left(\frac{1}{\text{١٠}} + \frac{(\text{س.} - \text{٣٩,٨})^2}{\text{١٦٨٣,٦}} \right)$$

وعند س = ٣٩,٨ = $\bar{\text{س}}$ فإن ع^2 (ص^) = ٢,١٣٢

$\therefore$ μ $_{\text{ص}^\wedge/\text{س.}}$ = ص^ $\pm$ ت $_{(\text{ن-٢ ، } \alpha/٢)}$ × ع (ص^)

= ٦٦,٤ $\pm$ ٢,٣٠٦ × $\sqrt{\text{٢,١٣٢}}$

= ٦٦,٤ + ٣,٣٦٧ = ٦٣,٠٣٣ & ٦٧,٧٦٧

مثال (٣٠) (نصف محلول)

البيانات التالية تمثل نتائج دراسة في نهاية سنة ما عن تكاليف الصيانة (ص) لآلات بيع المشروبات آليا بحسب العمر الزمني (س) وقد شملت العينة ١٠ ماكينات وكانت النتائج كما يلي : س = العمر الزمني للآلة ، ص : تكاليف الصيانة سنويا .

مجـ س = ٣٥ ، مجـ ص = ١٣٩٧ ، مجـ س ص = ٥٠٠٤

مجـ س٢ = ١٣٣ مجـ ص٢ = ١٩٦٦٦٥ ن = ١٠

<u>المطلوب :</u>

١- معادلة خط انحدار ص /س بطريقة المربعات الصغرى .

٢- قدر التكاليف المتوقعة لآلة ما بعد مرور عامين على استخدامها .

٣- استخدام تحليل التباين لاختبار β = صفر وحقق ما تصل إليه باستخدام اختبار ت .

٤- قدر β بفترة ثقة ٩٥% .

الحل:

١- ص = ١٠١,٦ + ١٠,٩ س

٢- ص^ عند س = ٢ هي : ص ^ = ١٢٣,٣ جنيه

٣- م.م.ك = ١٥٠٤,١ ، م.م.ب = ١٢٤٨,٠٥ ، ف = ٣٨,٩٩ ، ت = ٦,٢٤

٤- β = ١٠,٩ ± ٤,٠٢٧٩ = ٦,٩ & ١٤,٩

تمارين

(١) أ- ارسم الخط البياني المار بالنقط التالية :

(١) (١،١) & (٥ ، ٥)   (٢) (٠ ، ٣) ، (٣ ، ٠)

(٣) (-١ ، ١) & (٤، ٢)   (٤) (-٦ ، -٣) ، (٢ ، ٦)

ب- حدد كل من الميل والجزء المقطوع في الخطوط السابقة بيانيا .

جـ - أو جد المعادلات الرياضية التي تمر بالنقط في (أ)

(٢) ارسم الخطوط المستقيمة التالية بيانيا ثم حدد الميل والجزء المقطوع

(١) ص = س + ٤   (٢) ص = ٥- ٢س

(٣) ص = -٢س   (٤) ص = ٢ + ٦ س

(٣) أوجـد معادلة خط انحدار ص /س ، وكذلك معادلة خط انحدار س/ص في الحالات التالية مستخدما طريقة المربعات الصغرى :

(أ)

س	٧	٤	٦	٢	١	١	٣
ص	٢	٤	٢	٥	٧	٦	٥

(ب)

س	٨	٥	٤	٦	٢	٥	٣
ص	١	٣	-٦	٣	٧	٢	٥

(جـــ) ارســم الشــكل الإنتشاري لكل من (أ) ، (ب) ثم وفق عليها الخطوط التي توصلت إليها في (أ) ، (ب) .

(٤) بين أي المعادلات التالية تعبر عن معادلة انحدار بسيط وأيها تعبر عن انحدار متعدد .

(أ) ص = ٢ + ٣ س   (ب) ص = ٢+٣س$_{١}$ + ٥ س$_{٢}$

(ج) ص = ٧-٢س$_{١}$ + ٣ س$_{٢}$ (د) ص = - ٥ + ٣ س

(٥) ما هو الفرق بين الارتباط والانحدار ؟ وما هو المقصود بعبارة : انحدار ص على س ؟ حدد أيهما متغير مستقل وأيهما متغير تابع .

(٦) البيانات التالية تمثل عدد أيام الغياب (ص) ، وعدد سنوات العمل (س) في احدى الشركات ، لعينة عشوائية من ١٠ عمال بهذه الشركة

ص	٢	٠	٥	٦	٤	٩	٢	١٥	٠	٧
س	٧	٨	٢	٣	٥	٤	٦	١٠	٤	١١

المطلوب :

أ – رسم الشكل الإنتشاري لهذه البيانات .

ب– ايجاد معادلة خط انحدار ص / س بطريقة المربعات الصغرى .

جـ– التنبؤ بعدد أيام الغياب لعامل بالشركة يعمل بها منذ أربع سنوات.

د– احسب قيمة معامل التحديد ومعامل الارتباط .

هـ – اختبر الفرض القائل بعدم وجود علاقة بين (س ، ص) عند α = ٥ % مستخدما أسلوب تحليل التباين

(٧) البيانات التالية تمثل عدد سنوات الخدمة (س) وعدد العاملين الذين استقالوا من العمل باحدى الشركات (ص) :

عدد العاملين المستقيلين (ص)	٧	٦	٣	٢	٧
عدد سنوات الخدمة (س)	٤	٥	٦	٧	٣

المطلوب :

أ – معادلة خط الانحدار ص / س بطريقة المربعات الصغرى .

ب– معامل التحديد ومعامل الارتباط موضحا الفرق بينهما ثم معامل عدم التحديد .

جـ – اختبار معنوية معامل الانحدار مستخدما في ذلك كل من أسلوب فـترة الـثقة وأسلوب اختبار المعنوية (توزيع ت) وذلك عند مستوى معنوية ٥% .

(٨) ترغب احدى شركات التأمين على الحياة في تحديد العلاقة بين خبرة البائع وحجـم المبيعات له . سحبت عينة عشوائية من تسعة من مندوبي المبيعات حيـث سـجلت خبرة كل منهم بالسنوات (س) ومبيعاتهم السنوية (ص) في السنة الحالية بمئات الآلاف من الجنيهات فكانت

ص	١	٢	٣	٤	٥	٦	٧	٨	٩
س	٢	١	٣	٣	٤	٥	٦	٥	٧

المطلوب

I- تقدير المبيعات السنوية لمندوب خبرته ١٠ سنوات .

II- الخطأ المعياري لتقدير معادلة خط الإنحدار .

ج- اختـبار β = صـفر عند α = ٥% مستخدماً توزيع (ت) ثم دعم قرارك بأسلوب فترة الثقة .

د – أوجد فترة الثقة ٩٥% لمتوسط المبيعات عندما تكون :

(١) س = ٤,٥ ، (٢) س = ١٠

هـ – أوجد فترة الثقة ٩٥% للقيمة الفعلية ص عند س = ١٠ سنوات .

و – قارن بين النتائج التى حصلت عليها فى كل من (د) ، (هـ) ومعلقاً عليها .

(٩) تحـتفظ إحدى الشركات بسجل عن تكلفة صيانة كل ماكينة من ماكيناتها ، بالإضـافة إلـى عمر كل منها . لمعرفة العلاقة بين عمر الماكينة (س)

وتكلفة الصيانة (ص) ، سحبت عينة عشوائية من ٦ ماكينات وحصلنا على البيانات التالية :

الماكينة	١	٢	٣	٤	٥	٦
س	٢	١	٣	٢	١	٣
ص	٧٠	٤٠	١٠٠	٨٠	٣٠	١٠٠

المطلوب

I- معادلة خط انحدار ص /س بطريقة المربعات الصغرى .

II- تحديد تكلفة الصيانة لماكينة عمرها ٤ سنوات .

جـ- مستخدما اسلوب فترة الثقة ، اختبر β = صفر عند α = ١ %

د- ايجاد فترة الثقة ٩٩% لمتوسط التكلفة عند س = ٢ ثم عند س= ٤

هـ- ايجاد فترة الثقة ٩٩% للتكلفة الفعلية ص عند س = ٢

(١٠) اجريت تجربة لتحديد العلاقة بين كمية الأمطار بالبوصة (س) ومحصول القمح (ص) بالطن ، فحصلنا على البيانات التالية :

كمية الإمطار س	١	٢	٣	٤	٥	٥	٦	٧	٨	٩
كمية القمح ص	١	٣	٢	٥	٥	٤	٧	٦	٩	٨

المطلوب :

أ – معادلة خط انحدار ص / س بطريقة المربعات الصغرى .

ب- معامل التحديد ومعامل الارتباط ومعامل التحديد ومدلول كل منهم .

جـ- الخطأ المعياري لتقدير معادلة خط الانحدار .

د-اختبار الفرض القائل بعدم وجود علاقة بين كمية المطر ومحصول القمح (α=٥%) مستخدما في ذلك أساليب مختلفة من الاختبارات.

(١١) في عينة عشوائية حجمها ١٠ من طلبة احدى الجامعات ، جمعت بيانات عــن المعــدل التراكمي في المرحلة الثانوية (س) والمعدل التراكمي في الجامعة (ص) ، فحصلنا على المعلومات التالية :

مجـــ س = ٣٢,٢ مجـــ ص = ٣٠ مجـــ س ص =٩٨,٣٢

مجـــ س٢ = ١٠٥,٧٤ مجـــ ص٢ = ٩١,٩ ن= ١٠

المطلوب :

أ – معادلة خط انحدار ص/س .

ب – الخطأ المعياري لتقدير معادلة خط الانحدار .

جـــ– فترة الثقة ٩٥% لمعلمة الانحدار β .

د– اختــبار β = صفر مستخدما : (١) فترة الثقة (٢) اختبار المعنوية باستخدام توزيع (ت) (٣) تحليل التباين وذلك عند (α=٥%) .

هـــ– قدر متوسط ص بفترة ثقة ٩٥% عند (١) س= ٣٠،(٢) س= ٤٠

(١٢) أجريــت دراسة عن العلاقة بين كمية النتروجين في السماد (س) وكمية محصول القمح (ص) على عينة من ٢٠ فدان متماثلة ، فحصلنا على :

مجـــ س = ٢٣٠ مجـــ ص = ٢٦٠ مجـــ س ص = ٣٤٩٠

مجـــ س٢ = ٣١٤٤ مجـــ ص٢ = ٣٩٠٤

المطلوب :

I– معادلة خط الانحدار ص/ س بطريقة المربعات الصغرى

II– الخطأ المعياري لتقدير معادلة خط الانحدار .

جـــ – اختبار β = صفر مستخدما أسلوب تحليل التباين عند α = ٥%

د– قدر بفترة ثقة ٩٥% متوسط كمية القمح عند س = ١١,٥ ثم عند س = ١٥ ثم حدد أي التقديرين تفضل ولماذا ؟

(١٣) (أ) اذا كانت معادلة خط انحدار ص /س هي : ٥ص + ٣ س = ٤٢ ومعادلة خط انحدار س /ص هي : ٢ س+ ص = ٣٠ ، أوجد ما يلي:
١- معامل الارتباط بين س ، ص ٢- قيمة س عندما تكون ص = ١٢

(ب) إذا كان معامل الارتباط الخطي بين (س، ص) هو ٠,٨ ومعامل انحدار ص/س هو ١,٦ فإذا علمت أيضا أن $\bar{س}$ =١١ ، $\bar{ص}$ =٢٠ ، أوجد معامل انحدار س/ص وكذلك معادلتي خط انحدار ص/ س ، س /ص ثم قدر قيمة ص عندما تكون س =١٠ .

(١٤) (أ) إذا كانت معادلة خط انحدار ص / س هي : ص = ٠,١ س + ٧,٧ ومعادلة خط انحدار س / ص هي : س = ٥ ص – ٧ أوجد $\bar{س}$، $\bar{ص}$

(ب) إذا كان خطي الانحدار على الصورة التالية :

س + ٢ ص – ٥ = صفر & ٢ س + ٣ ص – ٨ = صفر

فإذا علمت أن تباين س اى ($ع^{٢}_{س}$) = ١٢ ، أوجد :

١- $\bar{س}$ ، $\bar{ص}$ ٢- $ع_{ص}$ ٣- معامل الارتباط بين س ، ص

(١٥) (أ) إذا كان معاملي خط الانحدار ص / س ، س / ص هما : ب = ٣,٢ ، د = ٠,٨ بين أن هناك خطأ في أحد هذه المعاملات .

(ب) إذا كان معاملي خط الانحدار ب =٠,٨ ، د = ٠,٢ فما هي قيمة معامل الارتباط

(ج) بفرض أنه فى معادلة خط الانحدار البسيط ، كانت ن = ١٥ ، $ر^{٢}$ = ٠,٨ فما هى قيمة وسيلة الاختبار ف إذا ما استخدم أسلوب تحليل التباين ، وما هى نتيجة اختبار β=صفر عند α = ٥% .

(د) بفرض أن معامل التحديد $ر^{٢}$ = ٠,٨ وأن الانحراف المعياري للظاهرة (ص) هو ٦ ، فما هو قيمة الخطأ المعياري لتقدير معادلة خط الانحدار ، $خ_{م}$ ؟

(١٦) (أ) استخدم البيانات التالية في ايجاد معادلتي خط انحدار ص/س ، س/ص ثم اعرضهما بيانيا موضحا مدلول نقطة التقاطع بينهما

مجـــ س= ٥٨٠ ، مجـــ ص =٣٧٠ ، مجـــ س ص= ١١٤٩٤

مجـــ $س^{٢}$ = ٤١٦٥٨ ، مجـــ $ص^{٢}$ = ١٧٢٠٦ ، ن = ١٢

(ب) في دراسة على عينة من ١٠ مفردات ، حصلنا على النتائج التالية:

ص = ٨ + ٠,٤ س ، $خ^{٢}_{م}$ = ١,٥ ، $ع^{٢}_{س}$ = ٤

المطلوب :

١- ميل خط انحدار س/ص ٢- معامل الارتباط بين (س ، ص)

(١٧) مـــن المعـــتاد في مجال التجارة إعطاء خصم عند شراء كمية كبيرة من السلعة . البيانات التالية توضح هذه الحقيقة .

سعر الوحدة (ص)	١٠	٨	٧	٦	٤
عدد الوحدات المشتراة (س)	١	٢	٣	٤	٥

المطلوب :

أ- عـــرض هـــذه البـــيانات بيانيا موضحا الجزء المقطوع وميل خط الإنحـــدار بعـــد أن تكـــون قد وفقت لها معادلة خط انحدار ص/س بطريقة المربعات الصغرى .

ب-اختـــبار β = صـــفر عند α = ٥% مستخدما اختبار ت ثم أعد الاختبار مرة أخرى باستخدام تحليل التباين . هل هناك اختلاف في القرار الذي تتخذه ؟ ولماذا ؟

جـ- قدر متوسط ص بفترة ثقة ٩٥% عند س = ٣ ثم عند س = ٥
ثم حدد أي الفترتين تفضل ولماذا ؟

(١٨) الجدول التالي يبين كمية المعروض من سلعة ما (ص) عند اسعار مختلفة (س) :

ص	١٢	١٤	١٠	١٣	١٧	١٢	١١	١٥
س	٥	١١	٧	٨	١١	٧	٦	٩

<u>المطلوب :</u>

أ- معادلة خط انحدار ص /س

ب-اختبر المعنوية الاحصائية لميل خط الانحدار β باستخدام توزيع ت ثم باستخدام فترة الثقة ، ثم دعم ذلك مستخدماً أسلوب تحليل التباين.

جـ - أوجد معامل التحديد ومعامل عدم التحديد ومدلول كل منهم .

د- قدر قيمة ص^ بفترة ثقة ٩٥% عند س = ١٠

<u>ملحوظة :</u>

عليك التأكد من الاجابات التالية :

أ- ص = ٦,٤٤ + ٠,٨٢ س

ب- ت = ب ÷ ع (ب) = ٠,٨٢ ÷ ٠,٢٥ = ٣,٢٨

جـ - ر٢ = ٠,٦٤٠٥

د- ص^ = ٦,٤٤ + ٠,٨٢ × ١٠ = ١٤,٦٤

$$\therefore \mu_{\text{ص}^\wedge/\text{س}} = ١٤,٦٤ \pm \text{ت}_{(٦ ، ٠,٠٢٥)} \times \text{ع}(\text{ص}^\wedge)$$

$$\text{حيث } \text{ع}^{٢}(\text{ص}^\wedge) = \frac{١٢,٩٤١٦}{٦}\left(١ + \frac{١}{٨} + \frac{(١٠-٨)^{٢}}{٣٤}\right) = ٢,٦٧$$

$$\therefore \mu_{\text{ص}^\wedge/\text{س}} = ١٤,٦٤ \pm ٢,٤٥ \times ١,٦٣ = ١٠,٦٥ \text{ ، } ١٨,٦٣$$

(١٩) الجـــدول التالى يوضح توزيع عينة من الأسر حسب حجم الدخل السنوى بالألف جنيه (س) وحجم الإنفاق بالألف جنيه (ص) والمطلوب :

١- معادلة خط انحدار ص/س .

٢- تقدير الإنفاق المتوقع لأسرة يبلغ دخلها السنوى ٦ (ألف جنيه).

س \ ص	-٣	-٥	-٧	-٩	المجموع
-٨	٣	٤	٢		٩
-١٠	٤	٦	٢		١٢
-١٢		٥	٣	٦	١٤
-١٤			٤	٣	٧
١٦-١٨			٥	٣	٨
المجموع	٧	١٥	١٦	١٢	٥٠

الفصل الرابع

تحليل الانحدار المتعدد

Multiple Regression Analysis

(١) مقدمة

(٢) فروض نموذج الانحدار .

(٣) تقدير معاملات الانحدار الجزئية .

(٤) معامل التحديد ومعامل التحديد المعدل .

(٥) الخطأ المعياري للتقدير .

(٦) معاملات الارتباط الجزئية .

(٧) الاستدلال الإحصائي عن معالم خط الانحدار .

(٨) اختبار المعنوية الكلية للانحدار : مدخل تحليل التباين .

(٩) اختبار مساهمة المتغيرات التفسيرية: مبدأ مجموع المربعات الإضافى .

(١٠) بعض مشاكل استخدام تحليل الانحدار :

١- مشكلة عدم ثبات التباين .

٢- مشكلة الازدواج الخطى .

٣- مشكلة الارتباط الذاتى للبواقى .

تمارين

الفصل الرابع

تحليل الانحدار المتعدد

Multiple Regression Analysis

(1) : مقدمة Introduction

فــي الانحدار الخطى البسيط كنا نهتم بدراسة تأثير متغير مستقل واحد س على المتغير التابع ص ، لكن من الناحية الواقعية والعملية فإن المتغير التابع دائمــاً يكــون واقعاً تحت تأثير عدة متغيرات مستقلة أو تفسيرية وليس متغير مستقل واحد . فمثلاً عند دراسة العلاقة بين الدخل والإنفاق ، كنا نفترض ضمنياً أن الدخــل (س) فقــط يؤثــر فى الإنفاق (ص) ، لكن النظرية الاقتصادية لهذه العلاقة نادرا ما تكون بهذه البساطة ، فبجانب الدخل ، هناك عددا من المتغيرات التفســيرية الأخــرى تؤثــر أيضاً فى الإنفاق مثل : ثروة المستهلك ، المستوى التعليمي ، المستوى الاجتماعى ...إلخ . مثال آخر : الطلب على سلعة ما (ص) لا يعــتمد فقــط على سعرها (س) ولكن يعتمد أيضا على أسعار السلع الأخرى ســواء المتنافســة أو المتكاملة معها ، ويعتمد أيضاً على دخل المستهلك وعلى الحالة الاجتماعية ...إلخ ، إلى غير ذلك من الأمثلة ، لذا فإننا نحتاج إلى تطوير وتوسيع نموذج الانحدار البسيط ليغطي عدداً أكبر من المتغيرات المستقلة ، وهذا يقودنــا إلى مناقشة نموذج الانحدار المتعدد ، أى النموذج الذي يضم العديد من المتغــيرات المســتقلة والتى تؤثر فى متغير تابع واحد (ص) . وسوف نكتفى بأبســط نمــوذج انحدار متعدد وهو النموذج الذي يضم متغير تابع واحد (ص) ومتغيرين مستقلين $س_1$ ، $س_2$ وصورته العامة فى مجتمع الدراسة هى :

$$ص = \beta + \beta_1 س_1 + \beta_2 س_2 + \psi \quad \ldots (1)$$

حيث :

β : مقدار ثابت أو هو قيمة ص عندما تكون $س_1 = س_2 =$ صفر

β_1 : معامل انحدار جزئي للمتغير ص على س١ مع اعتبار س٢ متغير ثابت أو تعرف β_1 بصورة أخرى على أنها مقدار التغير الذي يحدث فى التابع ص عـــندما يتغـــير المتغير المستقل س١ بوحدة واحدة مع فرض ثبات تأثير المتغير المستقل الآخر س٢ .

β_2 : معامل انحدار جزئي للمتغير ص على س٢ بفرض ثبات س١ أو هو مقدار التغير فى ص عندما تتغير س٢ بوحدة واحدة على فرض ثبات تأثير س١.

ψ : الحد العشوائي أو البواقى .

وعـــند التمثـــيل البيانى لمعادلة خط الانحدار المتعدد على ثلاث محاور مـــتعامدة فإننا نحصل على ما يسمي بمستوى الانحدار Regression plane والجـــزء المشـــترك المقطوع من المحاور الثلاث المتعامدة عبارة عن المقدار الثابت β .

(٣) فروض نموذج الانحدار : Assumptions of the Model

لا تخـــتلف هـــذه الفـــروض كثيراً عن فروض نموذج الانحدار البسيط ويمكن تلخيصها على النحو التالى :

١- توقع (ψ) = صفر .

٢- تباين (ψ) = σ^2 = مقدرا ثابت .

٣- تغاير الحدود العشوائية (ψ_j و ψ_i) = صفر .

٤- تغاير (س١ ، ψ) = تغاير (س٢ ، ψ) = صفر

٥- لا يوجد ارتباط ذاتي بين المتغيرات المستقلة أى تغاير (س١ ، س٢)= صفر

٦- التوزيع الاحتمالي للمتغير التابع ص هو التوزيع الطبيعي .

(٣) تقدير معاملات الانحدار الجزئية بطريقة المربعات الصغرى :

OLS Estimators

لتقدير المعالم المجهولة فى نموذج الانحدار المتعدد β ، β_1 ، β_2 فإننا نســـتخدم طريقة المربعات الصغرى والتى تعرضنا لها فى الباب السابق .وفق

هـذه الطريقة نقوم بسحب عينة عشوائية من مجتمع الدراسة ويسجل لكل مفردة فـيها ثـلاث قراءات تتعلق بالمتغيرات الثلاث ص ، $س_١$ ، $س_٢$ ويكون نموذج الانحدار بدلالة رموز العينة على النحو التالي :

$$ص = ب + ب_١ س_١ + ب_٢ س_٢ + خ \quad ...(٢)$$

حيـث خ هو حد البواقي أي العوامل العشوائية . وكما بينا من قبل فإن طريقة المربعات الصغرى تبحث في تدنيه مجموع مربعات البواقي ، مجـ $خ^٢$ إلى أقل قيمة ممكنة أي :

$$مجـ\ خ^٢ = مجـ\ (ص - ب - ب_١ س_١ - ب_٢ س_٢)^٢ \quad \text{أقل ما يمكن}$$

وعـن طـريق مفاضـلة مجـ $خ^٢$ جزئياً بالنسبة إلى ب ثم بالنسبة إلى $ب_١$ ثم بالنسـبة إلى $ب_٢$ ومساواة الناتج بالصفر ، نحصل على ثلاث معادلات وتسمي بالمعادلات الطبيعية وهذه المعادلات هي:

$$\left.\begin{array}{l} مجـ\ ص = ن\ ب + ب_١\ مجـ\ س_١ + ب_٢\ مجـ\ س_٢ \\ مجـ\ ص\ س_١ = ب\ مجـ\ س_١ + ب_١\ مجـ\ س_١^٢ + ب_٢\ مجـ\ س_١\ س_٢ \\ مجـ\ ص\ س_٢ = ب\ مجـ\ س_٢ + ب_١\ مجـ\ س_١\ س_٢ + ب_٢\ مجـ\ س_٢^٢ \end{array}\right\} ..(٣)$$

وبحـل تلـك المعادلات آنيا بأي طريقة جبرية ، نحصل على تقديرات المـربعات الصـغرى ب ، $ب_١$ ، $ب_٢$. جديـر بالذكر ان هناك طريقة تقليدية للوصـول إلـى تلـك المعـادلات الثلاث تتم على ثلاث مراحل : الأولى جمع المعادلـة (٢) ن مـن المرات والثانية ضرب طرفي المعادلة (٢) في المتغير المسـتقل $س_١$ ثـم جمعها والمرحلة الثالثة والأخيرة هي ضرب طرفي المعادلة الأصـلية (٢) فـي المتغير المستقل $س_٢$ ثم جمع الطرفين بذلك نصل إلى صيغ المعادلات الثلاث الطبيعية السابقة .

طريقة أخرى مبسطة لحساب معاملات الانحدار الجزئية :

إذا كانـت البيانات المتاحة عن المتغيرات ص ، $س_١$ ، $س_٢$ ذات أعداد كبيرة ويصعب التعامل معها مباشرة يدويا (كتربيع هذه الأعداد أو ضربها مثنى

مثنى) فإننا نلجأ إلى طريقة الانحرافات حيث تختزل تلك المتغيرات عن طريق طرح الأوساط الحسابية منها أى :

$ح_{ص} = ص - \bar{ص}$ ، $ح_{١} = س_{١} - \bar{س}_{١}$ ، $ح_{٢} = س_{٢} - \bar{س}_{٢}$

وعلى ذلك تختزل معادلة الانحدار المتعدد إلى الصور التالية :

$$ح_{ص} = ب_{١} ح_{١} + ب_{٢} ح_{٢} \quad ... (٤)$$

وبالتالى يمكن الحصول على الثوابت $ب_{١}$ ، $ب_{٢}$ من خلال حل المعادلتين :

$$\left. \begin{array}{l} مجـ\ ح_{ص} ح_{١} = ب_{١} مجـ\ ح^{٢}_{١} + ب_{٢} مجـ\ ح_{١} ح_{٢} \\ مجـ\ ح_{ص} ح_{٢} = ب_{١} مجـ\ ح_{١} ح_{٢} + ب_{٢} مجـ\ ح^{٢}_{٢} \end{array} \right\} ... (٥)$$

أما الثابت ب فتحصل عليه من المعادلة التالية :

$$ب = \bar{ص} - ب_{١} \bar{س}_{١} - ب_{٢} \bar{س}_{٢} \quad ... (٦)$$

مثال (١) :

استخدم البيانات الفرضية التالية فى إيجاد معادلة خط انحدار $ص/س_{١}$ ، $س_{٢}$ ثم قدر قيمة المتغير التابع ص عندما تكون $س_{١} = ٧$ ، $س_{٢} = ١٠$

ص	٣	١	٨	٣	٥
$س_{١}$	٣	١	٥	٢	٤
$س_{٢}$	٥	٤	٦	٤	٦

الحل :

ص	$س_{١}$	$س_{٢}$	$ص^{٢}$	$س^{٢}_{١}$	$س^{٢}_{٢}$	ص $س_{١}$	ص $س_{٢}$	$س_{١} س_{٢}$
٣	٣	٥	٩	٩	٢٥	٩	١٥	١٥
١	١	٤	١	١	١٦	١	٤	٤
٨	٥	٦	٦٤	٢٥	٣٦	٤٠	٤٨	٣٠
٣	٢	٤	٩	٤	١٦	٦	١٢	٨
٥	٤	٦	٢٥	١٦	٣٦	٢٠	٣٠	٢٤
٢٠	١٥	٢٥	١٠٨	٥٥	١٢٩	٧٦	١٠٩	٨١

بالتعويض بهذه المجاميع فى المعادلات الثلاث الطبيعية رقم (٣) :

٢٠ = ٥ ب + ١٥ ب$_{١}$ + ٢٥ ب$_{٢}$

٧٦ = ١٥ ب + ٥٥ ب$_{١}$ + ٨١ ب$_{٢}$

١٠٩ = ٢٥ ب + ٨١ ب$_{١}$ + ٥٥ ب$_{٢}$

وبحل تلك المعادلات بأى طريقة جبرية كالحذف او المحددات أو المصفوفات ، نحصل على التقديرات ب ، ب$_{١}$ ، ب$_{٢}$ على النحو التالى :

ب = ٤ ، ب$_{١}$ = ٢,٥ ، ب$_{٢}$ = -١,٥

وعلى ذلك تصبح معادلة خط الانحدار على الصورة :

ص = ٤ + ٢,٥ س$_{١}$ − ١,٥ س$_{٢}$

ولتقدير قيمة ص عند س$_{١}$ = ٧ ، س$_{٢}$ = ١٠

ص = ٤ + ٢,٥ × ٧ − ١,٥ × ١٠ = ٦,٥

(٤) معامل التحديد ومعامل التحديد المعدل R^2 and Adjusted R^2

معـــامل التحديد Coefficient of Determination هو ذلك المعامل الـــذي يبيـــن نســـبة التغير في المتغير التابع والتي يمكن تفسيرها بسب وجود المتغـــيرات المستقلة أو بمعنى آخر معامل التحديد يعطى نسبة تأثير المتغيرات المستقلة على المتغير التابع .

$$ر^{٢} = \frac{\text{التغير المفسر بسبب وجود المتغيرات المستقلة}}{\text{التغير الكلى فى التابع ص}} = \frac{\text{مجـ}(\hat{ص} - \bar{ص})^{٢}}{\text{مجـ}(ص - \bar{ص})^{٢}} \quad ... (٧)$$

حيـــث $\bar{ص}$: متوســـط قيم ص ، $\hat{ص}$: القيم التقديرية الناتجة من معادلة خط الانحدار المتعدد عند القيم المعطاة لــ س$_{١}$،س$_{٢}$ وهناك صيغة مبسطة لحساب ر^{٢}

$$ر^{٢} = \frac{ب_{١}\,\text{مجـ}(ص - \bar{ص})(س_{١} - \bar{س}_{١}) + ب_{٢}\,\text{مجـ}(ص - \bar{ص})(س_{٢} - \bar{س}_{٢})}{\text{مجـ}(ص - \bar{ص})^{٢}} \quad ...(٨)$$

$$= \frac{ب_{١}(\text{مجـ}\,ص\,س_{١} - \text{مجـ}\,ص \times \text{مجـ}\,س_{١}/ن) + ب_{٢}(\text{مجـ}\,ص\,س_{٢} - \text{مجـ}\,ص \times \text{مجـ}\,س_{٢}/ن)}{\text{مجـ}\,ص^{٢} - (\text{مجـ}\,ص)^{٢}/ن} \quad ... (٩)$$

وإذا اختزلنا المتغيرات بأوساطها الحسابية نصل إلى :

$$\text{ر}^{2} = \frac{\text{ب}_{1}\,\text{مجـ}\ \text{ح}_{\text{ص}}\,\text{ح}_{1} + \text{ب}_{2}\,\text{مجـ}\ \text{ح}_{\text{ص}}\,\text{ح}_{2}}{\text{مجـ}\ \text{ح}^{2}_{\text{ص}}} \qquad ...(١٠)$$

وتطبيقاً لذلك على المثال السابق وبواسطة المعادلة (٩) :

$$\text{ر}^{2} = \frac{٢,٥\ (٧٦ - ٢٠ \times ١٥/٥) - ١,٥\ (١٠٩ - ٢٠ \times ٢٥/٥)}{١٠٨ - (٢٠)^{2}/٥}$$

$$\text{ر}^{2} = \frac{٢,٥\ (١٦) - ١,٥\ (٩)}{٢٨} = \frac{٤٠ - ١٣,٥}{٢٨} = \frac{٢٦,٥}{٢٨} = ٠,٩٤٦$$

أى أن ٩٤,٦ % مــن التغــيرات الكلية فى التابع ص ترجع إلى وجود المتغــيرات المســتقلة س_{1} ، س_{2} أو بمعنى آخر المتغيرات المستقلة س_{1} ، س_{2} تؤثــر بنســبة ٩٤,٦% فـــى المتغير التابع أما الباقى ٥,٤% فترجع إلى تأثير العوامل العشوائية أى تأثير البواقى . أما الجذر التربيعي لمعامل التحديد فيعطي معامل الارتباط المتعدد ، أى أن ر = $\sqrt{٠,٩٤٦}$ = ٠,٩٧٣

يلاحظ أنه إذا كان هناك متغير مستقل ثالث س_{3} ، فإننا نضيف حد آخر فـــى بسط أى معادلة من معادلات معامل التحديد (٨ ، ٩ ، ١٠) يمثل أثر هذا المتغـــير علـــى العلاقة الكلية، وبالتالى تزداد قيمة معامل التحديد ، لأن البسط يزداد ويظل المقام ثابتاً لا يتغير ، فمثلاً إذا كان فى النموذج متغير مستقل س_{3} فإن ر^{2} تكون على الصورة التالية مثلاً :

$$\text{ر}^{2} = \frac{\text{ب}_{1}\,\text{مجـ}\,(\text{ص} - \bar{\text{ص}})(\text{س}_{1} - \bar{\text{س}}_{1}) + \text{ب}_{2}\,\text{مجـ}\,(\text{ص} - \bar{\text{ص}})(\text{س}_{2} - \bar{\text{س}}_{2}) + \text{ب}_{3}\,\text{مجـ}\,(\text{ص} - \bar{\text{ص}})(\text{س}_{3} - \bar{\text{س}}_{3})}{\text{مجـ}\,(\text{ص} - \bar{\text{ص}})^{2}}$$

معـــنى هذا أن معامل التحديد يتأثر بعدد المتغيرات المستقلة الداخلة فى تركيب نموذج الانحدار ، ويعد ذلك عيباً او نقصا فى قيمة معامل التحديد ، لأن

بعـــض هـــذه المتغـــيرات المستقلة قد يكون غير معنوى التأثير ، ولتلافى هذا القصور أو النقص يتعين علينا تصحيح أو تعديل معامل التحديد Adjusted R^2 وذلـــك بـــأن نـــأخذ فى الحسبان النقص الناشئ فى درجات الحرية عند إضافة متغيرات مستقلة جديدة .

وللتوضــــيح : فى الانحدار البسيط والذي يشمل متغير تابع ص ومتغير مســـتقل واحد س ، كانت درجات حرية البواقى (الخطأ العشوائي) هى (ن−٢) أما فى الانحدار المتعدد والذي يشـــمل متغير تابع ص ومتغيرين مستقلـــين ($س_١$ ، $س_٢$) فـــإن درجات حرية البواقى سنجدها (ن −٣) أى أنه بزيادة المتغيرات المســـتقلة تـــنخفض درجات حرية البواقى وعليه فمن المتوقع أن تنخفض قيمة معامل التحديد المعدل (يرمز له بالرمز $\bar{ر}^{٢}$) عن قيمته قبل التعديل :

$$\bar{ر}^{٢} = ١ - (١ - ر^{٢}) \times \frac{ن - ١}{ن - ك} \quad ...(١١)$$

حيث :

ن : حجم العينة ، ك: عدد معالم نموذج الانحدار بما فيها المقدار الثابت

$ر^{٢}$: معامل التحديد

<u>ويلاحظ على المعادلة (١١) ما يلى :</u>

١- عندما تكون ك = ١ فإن $\bar{ر}^{٢} = ر^{٢}$.

٢- عندما تكون ك > ١ فإن $\bar{ر}^{٢} < ر^{٢}$.

٣- عندما تكون ن كبيرة جدا ومع ثبات قيمة ك فإن $\bar{ر}^{٢} \simeq ر^{٢}$

٤- عـــندما تكون ن صغيرة جداً وفى نفس الوقت ك أكبر من ن فإن $\bar{ر}^{٢} < ر^{٢}$ وقـــد تكـــون ســـالبة القيمة أى قد يكون معامل التحديد المعدل سالب القيمة (لاحـــظ أن معـــامل الـــتحديد $ر^{٢}$ لا يمكن أن يكون سالب القيمة إذ أن حدة الأدنـــى هـــو الصـــفر) . فى هذه الحالة فإننا نعتبر معامل التحديد المعدل

مســاويا للصــفر .(تذكــر أن معامل التحديد ر^{٢} تتروواح قيمته بين الصفر والواحد الصحيح)

وبحساب معامل التحديد المعدل للمثال السابق نجد أن : ر^{٢} = ٠,٩٤٦ ، ن = ٥ ، ك = ٣ (وهــى ب ، ب$_{١}$ ، ب$_{٢}$). باســتخدام المعادلة (١١) نصل إلى معامل التحديد المعدل ر$^{/٢}$:

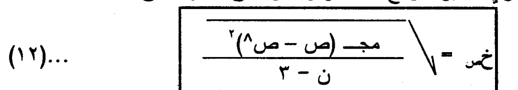
$$ر^{/٢} = ١ - (١ - ٠,٩٤٦) \times \frac{٥ - ١}{٥ - ٣} = ١ - ٠,٠٥٤ \times \frac{٤}{٢}$$

$$= ١ - ٠,١٠٨ = ٠,٨٩٢$$

ومن الواضح ان ر$^{/٢}$ أقل من ر^{٢}

(٥)الخطأ المعياري للتقدير Standard Error of the Regression Model

إذا كانــت صَ^ تمثل التقديرات المقابلة للقيم الفعلية ص ، والناتجة عن استخدام معادلة خط الانحدار المتعدد عند القيم المعطاة للمتغيرات المستقلة (س$_{١}$ ، س$_{٢}$) فــإن الفــرق بيــن القيم الفعلية ص والقيم التقديرية صَ^ ، يمثل الخطأ المعياري لتقدير نموذج الانحدار ويقاس على النحو التالى:



$$خ_ص = \sqrt{\frac{مجـ (ص - ص^)^{٢}}{ن - ٣}} \quad ...(١٢)$$

ويســمي المقــام (ن- ٣) بدرجــات الحــرية ، وهو عبارة عن حجم العينة ن مطروحاً منها عدد المعالم المطلوب تقديرها من بيانات العينة وهم ثلاثة : (ب ، ب$_{١}$ ، ب$_{٢}$) . وقد يستخدم البعض ن بدلاً من (ن -٣) بدافع السهولة ، لكن الدقة تستوجب استخدام (ن -٣) . وتقدير الخطأ المعياري وفق العلاقة (١٢) قد يمثل صعوبة حسابية خاصة وأن القيم التقديرية ص^ غالباً ما تكون قيم كسرية ، لذا فهــناك صــيغة أخــرى بديلة أكثر سهولة فى الاستخدام (تذكر أن مربع الخطأ المعياري للتقدير يسمي بتباين البواقى) هى :

$$\text{خ}_{\text{ص}} = \sqrt{(\text{مجـ ص}^{٢} - \text{ب مجـ ص} - \text{ب}_{١}\ \text{مجـ ص س}_{١} - \text{ب}_{٢}\ \text{مجـ ص س}_{٢}) \div (\text{ن}-٣)} \quad ...(١٣)$$

وتطبيقاً لذلك على بيانات المثال السابق نجد أن :

$$\text{خ}_{\text{ص}} = \sqrt{(١٠٨ - ٤ \times ٢٠ - ٢,٥ \times ٧٦ - (-١,٥) \times ١٠٩) \div (٥-٣)}$$

$$= \sqrt{(١٠٨ - ٨٠ - ١٩٠ + ١٦٣,٥) \div ٢} = \sqrt{(٢٧١,٥ - ٢٧٠) \div ٢}$$

$$= \sqrt{\frac{١,٥}{٢}} = \sqrt{٠,٧٥} = ٠,٨٦٦$$

تعليق :

يلاحـــظ أن قـــيمة الخطـــأ المعياري هنا لابد وأن تكون أقل من الخطأ المعياري ، لو كنا اكتفينا بمتغير مستقل واحد ، أى عند استخدام الانحدار البسيط ، ويـــرجع هـــذا النقص فى الخطأ المعياري للتقدير ، إلى إضافة متغير مستقل جديـــد إلـــى نموذج الانحدار البسيط . وعلى ذلك فالخطأ المعياري للتقدير فى نمـــوذج الانحدار المتعدد هو أقل دائماً من الخطأ المعياري فى نموذج الانحدار البسيط.

وكمـــا بيـــنا مـــن قـــبل فى حالة الانحدار البسيط ، العلاقة بين الخطأ المعـــياري للـــتقدير حول خط الانحدار ومعامل الارتباط البسيط ، فهناك أيضا علاقـــة مماثلـــة بيـــن الخطأ المعياري حول مستوى الانحدار ومعامل الارتباط المتعدد على الصورة :

$$\text{خ}^{٢}_{\text{ص}} = \frac{\text{ن}-١}{\text{ن}-٣} \times \text{ع}^{٢}_{\text{ص}} \left(١ - \text{ر}^{٢}_{\text{ص}(٢١)}\right) \quad ...(١٤)$$

ومن ناحية أخرى :

$$\text{ر}^{٢}_{\text{ص}(٢١)} = ١ - \frac{\text{ن}-٣}{\text{ن}-١} \times \frac{\text{خ}^{٢}_{\text{ص}}}{\text{ع}^{٢}_{\text{ص}}} \quad ...(١٥)$$

حيث:

$خ^2_{ص}$: مربع الخطأ المعياري للتقدير أو تباين البواقى Residuals .

$ع^2_{ص}$: تباين مفردات المتغير التابع $\bar{ص}$ حول متوسطها ص .

$ر^2_{ص(21)}$: مـــربع معـــامل الارتـــباط المتعدد بين ص وكل من $س_1$ ، $س_2$ أى معامل التحديد المتعدد (الأرقام ١ ، ٢ تعنى $س_1$ ، $س_2$) . والعلاقات (١٤) ، (١٥) هى علاقات إضافية لحساب الخطأ المعياري ومعامل الارتباط المتعدد كل بدلالـــة الآخـــر ، ويلاحظ أنه إذا كان حجم العينة ن أكبر من ٣٠ فمن الممكن إهمال المعامل $\frac{ن - ١}{ن - ٣}$

وتطبـــيقاً على ذلك ، وجدنا أن الخطأ المعياري للتقدير فى المثال السابق كان : $ح^2_{ص}$ = ٠,٧٥ فما هى قيمة معامل الإرتباط المتعدد ؟ لتطبيق المعادلة (١٥) ، يقتضى الأمر إيجاد تباين ص أى $ع^2_{ص}$ حيث :

$ع^2_{ص}$ =مجـــ(ص – $\bar{ص}$)^2÷(ن–١) = [مجـــ $ص^2$–(مجـــ ص)^2/ن] ÷ (ن–١)

ومن بيانات المثال (١) :

$ع^2_{ص}$ = [١٠٨ – (٢٠)^2/٥] ÷ ٤ = ٢٨ ÷ ٤ = ٧ ، ومن المعادلة (١٥) :

$$ر^2_{ص(21)} = ١ - \frac{٥-٣}{٥-١} \times \frac{٠,٧٥}{٧} = ١ - ٠,٠٥٤ = ٠,٩٤٦$$

أما معامل الارتباط المتعدد =$\sqrt{٠,٩٤٦}$ = ٠,٩٧٣ وهذه النتائج سبق الحصول علـــيها عـــند الحديث عن معامل التحديد . من ناحية أخرى إذا فرض و علمنا معـــامل الارتـــباط المتعدد ، فإنه يمكن استخدام المعادلة (١٤) في إيجاد الخطأ المعياري للتقدير .

(٦) معاملات الارتباط الجزئية : Partial Correlation Coefficients

يقيس معامل الارتباط الجزئي قوة العلاقة الإرتباطية بين المتغير التابع ص و أحـــد المتغـــيرات المستقلة بعد حذف تأثر – أو مع تثبيت – المتغيرات المســـتقلة الأخرى . فمثلا $ر_{ص1.2}$ تعبر عن الارتباط الجزئى بين المتغير التابع

ص و المتغير المستقل $س_1$ مع ثبات تأثير المتغير المستقل الثانى $س_2$ و يعرف $ر_{ص2.1}$ على نفس النمط . وتتراوح قيمة معامل الارتباط الجزئى بين −1 ، +1 (كما هو الحال بالنسبة لمعاملات الارتباط البسيطة) . وتقاس معاملات الارتباط الجزئــية إما بدلالة معامل الارتباط المتعدد أو بدلالة معاملات الارتباط الخطية البسيطة على النحو التالى :

$$ر_{ص1.2} = \sqrt{1 - \frac{1 - ر^2_{ص(21)}}{1 - ر^2_{ص2}}} \quad ...(16)$$

$$ر_{ص2.1} = \sqrt{1 - \frac{1 - ر^2_{ص(21)}}{1 - ر^2_{ص1}}} \quad ...(17)$$

حيث $ر^2_{ص(21)}$: مربع معامل الارتباط المتعدد

$ر^2_{ص1}$ ، $ر^2_{ص2}$ مربع معامل الارتباط البسط بين ص ، $س_1$ وكذلك بين ص ، $س_2$. (الأرقام 1 ، 2 تعنى المتغيرات $س_1$ ، $س_2$) فمثلا

$$ر_{ص1} = \frac{مجـ ص س_1 - مجـ ص \times مجـ س_1 / ن}{\sqrt{[مجـ ص^2 - (مجـ ص)^2 / ن] [مجـ س_1^2 - (مجـ س_1)^2 / ن]}}$$

وبنفس الطريقة يمكن ايجاد المعاملات $ر_{ص2}$ ، $ر_{21}$.

مــن ناحــية أخــرى ، يمكن قياس الارتباط الجزئى بدلالة معاملات الارتباط البسيطة على النحو التالى :

$$ر_{ص1.2} = \frac{ر_{ص1} - ر_{ص2} \; ر_{21}}{\sqrt{1 - ر^2_{ص2}} \sqrt{1 - ر^2_{21}}} \quad ...(18)$$

$$ر_{ص2.1} = \frac{ر_{ص2} - ر_{ص1} \; ر_{21}}{\sqrt{1 - ر^2_{ص1}} \sqrt{1 - ر^2_{21}}} \quad ...(19)$$

بالطــبع لابد وأن تختلف قيمة معامل الارتباط الجزئى عن قيمة معامل الارتباط البسيط ، وذلك بسبب إدخال متغيرات مستقلة جديدة فى نموذج الانحدار . جدير

بالذكـر أن معـاملات الارتـباط الجزئـية تستخدم فى تحديد الأهمية النسبية للمتغيرات المفسرة (المستقلة) المختلفة فى نموذج الانحدار المتعدد ، من ناحية أخـرى ، يمكـن استنتاج علاقة أخرى أضافية لمعامل الارتباط المتعدد بدلالة معاملات الارتباط الخطية البسيطة وهى على الصورة .

$$\text{ر}_{\text{ص}(٢١)} = \sqrt{[\text{ر}^٢_{\text{ص}١} + \text{ر}^٢_{\text{ص}٢} - ٢ \times \text{ر}_{\text{ص}١} \times \text{ر}_{\text{ص}٢} \times \text{ر}_{٢١}] \div [١ - \text{ر}^٢_{٢١}]} \quad ...(٢٠)$$

مثال (٣)

<u>أولا :</u> بفرض أنك حصلت على البيانات التالية :

$\text{ر}_{\text{ص}١} = ٠,٧$ ، $\text{ر}_{\text{ص}٢} = ٠,٨$ ، $\text{ر}_{٢١} = ٠,٩$

أوجد معاملات الارتباط الجزئية و الكلية

<u>ثانيا :</u> بفرض أنك حصلت على البيانات التالية :

$\text{ر}_{\text{ص}(٢١)} = ٠,٩$ ، $\text{ر}_{\text{ص}١} = ٠,٧$ ، $\text{ر}_{\text{ص}٢} = ٠,٦$

أوجد معاملات الارتباط الجزئية .

الحل :

<u>أولاً :</u> $\text{ر}_{\text{ص}١} = ٠,٧$ ، $\text{ر}_{\text{ص}٢} = ٠,٨$ ، $\text{ر}_{٢١} = ٠,٩$

بالتعويض فى المعادلات (١٨) ، (١٩)

$$\text{ر}_{\text{ص}١.٢} = \frac{\text{ر}_{\text{ص}١} - \text{ر}_{\text{ص}٢}\ \text{ر}_{٢١}}{\sqrt{١ - \text{ر}^٢_{\text{ص}٢}}\ \sqrt{١ - \text{ر}^٢_{٢١}}} = \frac{٠,٧ - ٠,٨ \times ٠,٩}{\sqrt{١ - ٠,٦٤}\ \sqrt{١ - ٠,٨١}}$$

$$= \frac{٠,٧ - ٠,٧٢}{\sqrt{٠,٣٦}\ \sqrt{٠,١٩}} = \frac{-٠,٠٢}{٠,٢٦٦١٥} = -٠,٠٧٦$$

$$\text{ر}_{\text{ص}٢.١} = \frac{\text{ر}_{\text{ص}٢} - \text{ر}_{\text{ص}١}\ \text{ر}_{٢١}}{\sqrt{١ - \text{ر}^٢_{\text{ص}١}}\ \sqrt{١ - \text{ر}^٢_{٢١}}} = \frac{٠,٨ - ٠,٧ \times ٠,٩}{\sqrt{١ - ٠,٤٩}\ \sqrt{١ - ٠,٨١}}$$

$$= \frac{٠,١٧}{٠,٣١١٣} = ٠,٥٤٦$$

أما معامل الارتباط المتعدد فهو على الصورة رقم (٢٠)

$$\text{ر}_{\text{ص}(٢١)} = \sqrt{[\text{ر}^{٢}_{\text{ص}١} + \text{ر}^{٢}_{\text{ص}٢} - ٢ \times \text{ر}_{\text{ص}١} \times \text{ر}_{\text{ص}٢} \times \text{ر}_{٢١}] \div [١ - \text{ر}^{٢}_{٢١}]}$$

$$\text{ر}_{\text{ص}(٢١)} = \sqrt{[٠,٤٩ + ٠,٦٤ - ٢ \times ٠,٧ \times ٠,٨ \times ٠,٩] \div [١ - ٠,٨١]}$$

$$= \sqrt{(١,١٣ - ١,٠٠٨) \div ٠,١٩} = \sqrt{٠,١٢٢ \div ٠,١٩} = ٠,٦٤$$

<u>ثانيا :</u> $\text{ر}_{\text{ص}(٢١)} = ٠,٩$ ، $\text{ر}_{\text{ص}١} = ٠,٧$ ، $\text{ر}_{\text{ص}٢} = ٠,٦$

بالتعويض فى المعادلات (١٦) ، (١٧) :

$$\text{ر}_{\text{ص}٢.١} = \sqrt{١ - \frac{١ - \text{ر}^{٢}_{\text{ص}(٢١)}}{١ - \text{ر}^{٢}_{\text{ص}٢}}} = \sqrt{١ - \frac{١ - ٠,٨١}{١ - ٠,٣٦}} = \sqrt{١ - \frac{٠,١٩}{٠,٦٤}}$$

$$= \sqrt{١ - ٠,٢٩٧} = \sqrt{٠,٧٠٣} = ٠,٨٣٨$$

$$\text{ر}_{\text{ص}١.٢} = \sqrt{١ - \frac{١ - \text{ر}^{٢}_{\text{ص}(٢١)}}{١ - \text{ر}^{٢}_{\text{ص}١}}} = \sqrt{١ - \frac{١ - ٠,٨١}{١ - ٠,٤٩}} = \sqrt{١ - \frac{٠,١٩}{٠,٥١}}$$

$$= \sqrt{١ - ٠,٣٧٣} = \sqrt{٠,٦٢٧} = ٠,٧٩٢$$

(٧) الاستدلال الإحصائى عن معالم خط الإنحدار

Inferences About The β Parameters

بيـــنا من قبل أن $\beta_{١}$ تقيس مقدار التغير فى التابع ص عندما تتغير س١ بوحدة واحدة مع ثبات تأثير المتغير المستقل الآخر س٢ ، كذلك تبين $\beta_{٢}$ مقدار التغـــير فى التابع ص الناشئ عن التغير فى س٢ بوحدة واحدة مع ثبات س١ . لكـــن إلى أى مدى يمكن الاعتماد على تقديرات تلك المعالم أى على ب١ ، ب٢ فى التنبؤ بقيمة المتغير التابع ص؟ الإجابة تكمن فى عملية الإستدلال الإحصائى حول تلك المعالم أى حول $\beta_{١}$ ، $\beta_{٢}$ ، (الاستدلال عن العلمة β –المقدار الثابت–

ليست موضع اهتمام ولا يدخل تقديرها فى أى عمليات حسابية قد نحتاج إليها) .
ويقصـــد بالاستدلال الإحصائى عن تلك المعالم هو : (١) إنشاء فترات ثقة لتلك المعـــالم (٢) إختبارات المعنوية لكل معلمة من تلك المعالم . لكن هذا الإستدلال يقتضى معرفة توزيع المعاينة لتقديرات تلك المعالم ، أى توزيع المعاينة لكل من $ب_١$ ، $ب_٢$ ومـــنها يمكن استنتاج الخطأ المعيارى لتلك التقديرات . عموما ودون الدخـــول فـــى تفصيلات رياضية ، ومع فرض تحقق شروط أو فروض نموذج الإنحدار المتعدد ، فإن تباين تقديرات طريقة المربعات الصغرى $ب_١$ ، $ب_٢$ على الصورة التالية :

$$ع^٢ (ب_١) = خ^٢_{ص} \times \frac{مجـ (س_٢ - \bar{س}_٢)^٢}{م} \quad ...(٢١)$$

$$ع^٢ (ب_٢) = خ^٢_{ص} \times \frac{مجـ (س_١ - \bar{س}_١)^٢}{م} \quad ...(٢٢)$$

حيث :

$خ^٢_{ص} = مجـ(ص - \hat{ص})^٢ \div (ن - ٣)$ أو الصورة المبسطة لها بالمعادلة رقم (١٣)

$م = مجـ(س_١ - \bar{س}_١)^٢ \times مجـ(س_٢ - \bar{س}_٢)^٢ - [مجـ (س_١ - \bar{س}_١)(س_٢ - \bar{س}_٢)]^٢$

أما ع ($ب_١$) ، ع ($ب_٢$) فهي الأخطاء المعيارية لتلك التقديرات .

وعـــن طريق معرفة الأخطاء المعيارية ع($ب_١$) ، ع ($ب_٢$) يمكن إنشاء فـــترات الثقة للمعالم $\beta_١$ ، $\beta_٢$ و جراء اختبارات المعنوية لها فمثلا ، فترة الثقة للمعلمة $\beta_١$ تكون على الصورة

$$\beta_١ = ب_١ \pm ت_{(ن-٣ ، \alpha/٢)} \times ع (ب_١) \quad ... (٢٣)$$

بالمثل :

$$\beta_٢ = ب_٢ \pm ت_{(ن-٣ ، \alpha/٢)} \times ع (ب_٢) \quad ... (٢٤)$$

أما فيما يتعلق باختبار المعنوية للمعلمة β_1 فتتم على النحو التالي :

١- الفرض العدمى : β_1 = صفر (لا يوجد تأثير للمتغير المستقل $س_1$ على التابع ص)

٢-الفرض البديل : $\beta_1 \neq$ صفر يوجد تأثير للمتغير $س_1$ على التابع ص

$$٣- ت = \frac{ب_1 - \beta_1}{ع\,(ب_1)} = \frac{ب_1}{ع(ب_1)}$$ ، حيث β_1 = صفر

٤- ت الجدولية : $ت_{ن-٣\,،\,\alpha/٢}$

٥- المقارنة بين ت المحسوبة ، ت الجدولية .

٦- اتخاذ القرار بقبول أو رفض الفرض العدمى .

ونتبع نفس الخطوات عند إختبار β_2 = صفر .

مثال (٣)

مستخدما بيانات مثال (١) ، المطلوب

(أ) فترة ثقة للمعلمة β_2 عند α = ٥%

(ب) اختبار معنوية المعلمة β_1 عند α = ٥%

الحل

(أ) فترة الثقة ٩٥% للمعلمة β_2 :

$$\beta_2 = ب_2 \pm ت_{(ن-٣\,،\,\alpha/٢)} \times ع(ب_2)$$

حيث $ب_2$ = −١,٥ ، $ت_{(٥-٣،\,٠,٠٢٥)} = ت_{٢،\,٠,٠٢٥}$ = ٤,٣٠٣

أما الخطأ المعيارى ع ($ب_2$) ووفقا للمعادلة (٢٢) :-

$$ع(ب_2) = خ_م \times \sqrt{\frac{مجـ\,س_1^2 - (مجـ\,س_1)^2 / ن}{م}}$$

حيث $خ_م$ سبق الحصول عليها = ٠,٨٦٦ أما (م) فهى:

$$م = [مجـ\,س_1^2 - (مجـ\,س_1)^2 / ن] \times [مجـ\,س_2^2 - (مجـ\,س_2)^2 / ن]$$

– [مجـ س١س٢ – مجـ س١ × مجـ س٢ / ن]٢

م = [٥٥– (١٥)٢/٥]×[١٢٩ – (٢٥)٢/٥] – [٨١ – ١٥ ×٢٥/٥]٢

= [١٠ × ٤] – [٦]٢ = ٤٠ – ٣٦ = ٤

∴ ع(ب٢) = ٠,٨٦٦ × $\sqrt{\frac{١٠}{٤}}$ = ٠,٨٦٦ × ١,٥٨١ = ١,٣٦٩

∴ β_2 = –١,٥ ± ٤,٣٠٣ × ١,٣٦٩ = –١,٥٠ ± ٥,٨٩

= + ٤,٣٩ ، – ٧,٣٩

(ب) إختبار معنوية β_1 = صفر

١– الفرض العدمى β_1 = صفر

٢– الفرض البديل β_1 ≠ صفر

٣– ت = $\frac{ب_1 - \beta_1}{ع\,(ب_1)}$ = $\frac{ب_1}{ع(ب_1)}$

حيث ب١ = ٢,٥ أما ع (ب١) فهى

ع(ب١) = خـ × $\sqrt{\frac{مجـ\ س_2^2 - (مجـ\ س_2)^2/ن}{م}}$ = ٠,٨٦٦× $\frac{٤}{٤}$ = ٠,٨٦٦

∴ ت = $\frac{٢,٥}{٠,٨٦٦}$ = ٢,٨٨

٤– ت الجدولية = ت (ن–٣ ، α /٢) = ت (٢ ، ٠,٠٢٥) = ٤,٣٠٣

٥ – المقارنة ت المحسوبة أقل من ت الجدولية

٦– القرار : قبول الفرض العدمى

∴ β_1 = صفر ، أى لا يوجد تأثير معنوى للمتغير س١ على التابع ص

ملحوظة :

مـن الممكـن أن تسـتخدم فـترة الـثقة للمعملة β_1 أو β_2 فى إختبار β_1 = صفر أو β_2 = صفر وذلك بالبحث عن وجود أو عدم وجود الصفر داخل

حـــدى الـــثقة . فـــإذا وقع الصفر داخل حدى الثقة ، لكان القرار قبول الفرض العدمـــى بـــأن β_1 = صــــفر مثلا أما إذا و قع الصفر خارج حدى الثقة ، لكان القرار هو رفض الفرض العدمى أى قبول الفرض البديل بإن $\beta_1 \neq$صفر.

(٨) إختبار المعنوية الكلية للإنحدار: مدخل تحليل التباين

Analysis Of Variance

أحـــد أهداف تحليل نموذج الانحدار ، هو تحديد ما إذا كانت المتغيرات التفســـيرية الداخلة فى النموذج ككل أى مجتمعة مع بعض ، ذات علاقة إرتباط حقيقـــية بالمتغير التابع ص أم لا . لنفرض أنه لا توجد علاقة حقيقية بين التابع ص و المتغـــيرات التفســـيرية أو المســـتقلة س١ ، س٢. هذا الفرض يعنى أن معـــاملات الانحـــدار فى مجتمع الدراسة β_1 ، β_2 = صفر . مقابل ذلك يظهر الفرض البديل الذى يعنى أن هناك أحد المتغيرات المستقلة ــ أن لم يكن كلها ــ على الأقل ذو علاقة حقيقية بالمتغير التابع ص أى $\beta_1 \neq \beta_2 \neq$ صفر .

بـــنفس الأسلوب الذى اتبع فى حالة استخدام تحليل التباين فى الانحدار البسيط ، فإننا نقسم الإختلافات الكلية فى التابع ص إلى جزئين (١) تغيرات أو اخـــتلافات ناتجـــة عن وجود المتغيرات التفسيرية س١ ، س٢ (٢) تغيرات أو اختلافات ناتجة عن عوامل عشوائية أو ما تسمى أحيانا بالبواقى . أى أن :

مجمـــو ع مربعات الاختلافات الكلية = مجموع المربعات المفسر بوجود المتغيرات المستقلة + مجموع المربعات المفسر (البواقى)

وعندما تقترن مجموع المربعات بدرجات الحرية ، نحصل على التباين ، أو متوســـط مجموع المربعات . وعن طريق قسمة متوسط مجموع المربعات المفسر على متوسط مجموع المربعات غير المفسر (البواقى) نحصل على متغير عشوائى (ف) ، وهو متغير يتبع توزيع ف بدرجات حرية (ك-١) ، (ن-ك) ، حيث ك :عدد المعالم المقدرة . بمقارنة قيمة ف مع قيمة جدولية من توزيع ف عـــند درجتى حرية (ك -١) ، (ن-ك) ، α يمكن أى نصل إلى قرار بقبول أو

رفض الفرض $\beta_1 = \beta_2 =$ صفر .

وبترتيــب مجاميع المربعات السابقة داخل جدول تحليل التباين ـــ بنفس الأسلوب الذى اتبع فى حالة الانحدار البسيط ـــ مقرونة بدرجات الحرية ، نصل إلى الأحصاء (ف) على الصورة التالية:

$$\text{ف} = \frac{\text{مجموع المربعات المفسر} \div (\text{ك} - 1)}{\text{مجموع المربعات غير المفسر} \div (\text{ن} - \text{ك})}$$

$$= \frac{[\text{ب}_1 (\text{مجـ ص س}_1 - \text{مجـ ص} \times \text{مجـ س}_1/\text{ن}) + \text{ب}_2 (\text{مجـ ص س}_2 - \text{مجـ ص} \times \text{مجـ س}_2/\text{ن})] \div (\text{ك} - 1)}{\text{مجـ خ}^2_{\text{س}} = \text{مجـ} (\text{ص} - \hat{\text{ص}})^2 \div (\text{ن} - \text{ك})}$$

حيث ك : عدد المعالم المقدرة : ب ، ب، ب٢ ، ن : حجم العينة

جديــر بالذكر أن هناك صورة رياضية بديلة لصورة ف السابقة ، وهي تعتمد على معلومية معامل التحديد المتعدد ر^2 ، ومن السهل الوصول إليها وهي على الصورة :

$$\text{ف} = \frac{\text{ر}^2 \div (\text{ك} - 1)}{(1 - \text{ر}^2) \div (\text{ن} - \text{ك})} = \boxed{\frac{\text{ر}^2}{1 - \text{ر}^2} \times \frac{\text{ن} - \text{ك}}{\text{ك} - 1}} \quad ...(٢٥)$$

مثال (٤):

مستخدما بيانات مثال (١)

المطلـــوب اختــبار $\beta_1 = \beta_2 =$ صـــفر عند $\alpha = ٥\%$ ، مستخدما في ذلك أسلوب تحليل التباين .

الحل :

يقتضي أسلوب تحليل التباين حساب المكونات التالية :

مجموع المربعات الكلي في المتغير التابع ص =

$$\text{م.م.ك} = \text{مجـ ص}^2 - (\text{مجـ ص})^2/\text{ن}$$

$$= ١٠٨ - (٢٠)^2 / ٥ = ١٠٨ - ٨٠ = ٢٨$$

مجموع المربعات المفسر بوجود $س_1$ ، $س_2$

= $ب_1$[مجـ ص $س_1$ – مجـ ص×مجـ $س_1$/ ن]

+ $ب_2$ [مجـ ص $س_2$–مجـ ص×مجـ $س_2$ / ن]

م.م. بسبب الانحدار = ٢,٥ [٧٦ –٢٠ × ١٥/٥] – ١,٥ [١٠٩– ٢٠ × ٢٥/٥]

= ٢,٥ × ١٦ – ١,٥ × ٩ = ٤٠–١٣,٥ = ٢٦,٥

جدول تحليل التباين

المصدر	م.م	د . ح	م.م.م	ف
م.م. بسبب انحدار ص/$س_1$ ، $س_2$	٢٦,٥	ك–١ = ٢	١٣,٢٥	١٧,٦٦٧
م.م.د (البواقي)	١,٥	ن–ك = ٢	٠,٧٥ = $خ^2_س$	
م.م.ك	٢٨	ن–١ =٤		

خطوات الاختبار :

١– الفــرض العدمى : $\beta_1 = \beta_2 =$صفر (لا توجد علاقة بين ص وكل من $س_1$،$س_2$)

٢– الفــرض الــبديل : $\beta_1 \neq \beta_2 \neq$صــفر (توجد علاقة خطية من المتغيرات الثلاث)

٣– ف = ١٧,٦٦٧

٤– ف الجدولية : ف (٢،٢،٥%) = ١٩

٥– المقارنة : ف المحسوبة أقل من ف الجدولية .

٦– القرار : قبول الفرض العدمي .

∴ $\beta_1 = \beta_2 =$ صفر أي لا توجد علاقة حقيقية بين ص وكل من $س_1$ ، $س_2$ ومــن ثم فإن هذه العلاقة لا تصلح في عملية التنبؤ بالمتغير التابع ص .

ملحوظة :

من الممكن استخدام معامل التحديد ر ٢ ومعامل عدم التحديد (١–ر٢) في قياس قيمة ف باستخدام العلاقة (٢٥) على النحو التالي .

$$ف = \frac{ر^٢}{١-ر^٢} \times \frac{ن-ك}{ك-١}$$ ، حيث ك = ٣ ، ن = ٥ .

فـــإذا كانت ر ٢ معلومة من بيانات سابقة [ر ٢ = ٢٦,٥ ÷ ٢٨ = ٠,٩٤٦٤] ، فإن قيمة ف تصبح :

$$ف = \frac{٠,٩٤٦٤}{١-٠,٩٤٦٤} \times \frac{٥-٣}{٣-١} = \frac{٠,٩٤٦٤}{٠,٠٥٣٦} \times \frac{٢}{٢} = ١٧,٦٦٧$$

وهي نفس القيمة التي ظهرت في جدول تحليل التباين .

(٩) اختبار مساهمة المتغيرات التفسيرية : مبدأ مجموع المربعات الإضافي
The Extra Sum of Squares Principle

بيـــنا من قبل كيف يمكن الاعتماد على أسلوب تحليل التباين في اختبار العلاقـــة الكلـــية لـــنموذج الانحدار ، أي اختبار مدى وجود علاقة حقيقية بين المتغيرات التفسيرية ككل ($س_١$،$س_٢$) مع المتغير التابع ص أي اختبار $\beta_١ = \beta_٢$ = صـــفر . أيضا بينا من قبل كيف يتم اختبار مدى وجود علاقة حقيقية بين كل متغير تفسيري على حدة $س_١$ ، $س_٢$ والمتغير التابع ص ، أي اختبار $\beta_١$= صفر واختـــبار $\beta_٢$ = صفر . في اختبار العلاقة الإجمالية ، استخدم توزيع ف ، وفي اختـــبار كل متغير على حدة ، استخدم توزيع ت . غير ان هناك وسيلة أخرى لاختـــبار معنوية $\beta_١$ أو اختبار معنوية $\beta_٢$ وذلك بالبحث في مدى أهمية وجود المتغـــير $س_١$ أو المتغير $س_٢$ في نموذج الانحدار ، بمعنى ، هل وجود المتغير $س_١$ له أهمـــية عـــند دراسة العلاقة بين ص ، $س_٢$ أم يمكن إهماله وحذفه من نموذج الانحدار ، أيضا ، هل من المفيد والضروري أن يظهر المتغير $س_٢$ في نمـــوذج الانحدار أم يمكن إهماله وحذفه ؟ هذه الأسئلة سبق أن أجيب عنها عند

اختـــبار معنوية معاملات خط الانحدار باستخدام اختبار ت . لكن هناك وسيلة أخـــرى يمكن الاستعانة بها لاختبار معنوية المتغيرات التفسيرية كل على حدة ، تعتمد على أسلوب تحليل التباين والفكرة الأساسية هنا يمكن إيضاحها كالآتي :

(١) إذا كـــنا نبحـــث في مدى أهمية وجود المتغير المستقل س$_{١}$ مثلا في نموذج الانحـــدار ، فإنـــنا نقـــارن بيـــن مجموع المربعات بسبب وجود س$_{١}$،س$_{٢}$ معاً ومجموع المربعات بسبب وجود س$_{٢}$ ، والفرق بين المجموعين لا بد وأن يرجع إلـــى مجموع المربعات الإضافي Extra sum of squares للمتغير س$_{١}$ أي الـــزيادة في مجموع المربعات بسبب وجود س$_{١}$. أي أننا نقارن بين نموذجين للانحدار ، الأول بدلالة س$_{١}$،س$_{٢}$ والثاني بدلالة س$_{٢}$ فقط أي :

النموذج الأول : ص = ب + ب$_{١}$س$_{١}$ + س$_{٢}$ س$_{٢}$

النموذج الثانى : ص = ب + ب$_{٢}$ س $_{٢}$

أما مجموع المربعات في ظل النموذج الأول والنموذج الثاني فهي على الصورة

م.م. بسبب انحدار ص/س$_{١}$ س$_{٢}$= ب$_{١}$(مجـ ص س$_{١}$ −مجـ ص×مجـ س$_{١}$/ن)

+ ب$_{٢}$(مجـ ص س$_{٢}$−مجـ ص×مجـ س$_{٢}$/ن)

م.م. بسبب انحدار ص/س$_{٢}$ = (مجـ ص س$_{٢}$ − مجـ ص × مجـ س$_{٢}$/ ن)^{٢} / مجـ س^{٢} $_{٢}$ − (مجـ س$_{٢}$)^{٢} / ن ...(٢٦)

وإذا كـــان م.م بســـبب انحـــدار ص / س$_{١}$ س$_{٢}$ له (٢) درجة حرية ، فإن م.م. بسبب انحدار ص / س$_{٢}$ ، يكون له درجة حرية واحدة .

وبترتيب تلك المجاميع مع مجموع المربعات الكلي ومجموع مربعات البواقي أو الخطـــأ العشوائي في جدول تحليل التباين ، يمكن اختبار مدى أهمية وجود س$_{١}$ فـــي نموذج الانحدار أي اختبار معنوية $\beta_{١}$ = صفر وذلك بإيجاد قيمة الإحصاء ف حيث :

ف = مجموع المربعات الإضافي أي مساهمة س$_{١}$ / تباين البواقي = غ^{٢} $_{س}$ ...(٢٧)

وبمقارنـــة قـــيمة ف السابقة مع قيمة جدولية من توزيع ف بدرجات حرية ١ ، (ن – ك) يمكــن اتخاذ قرار بقبول أو رفض الفرض العدمي بأن β_1 = صفر . بالطبع وكما هو متوقع لا بد وأن تتفق نتيجة القرار هنا مع القرار الذي يمكن أن نصـــل إليه لو استخدمنا اختبار (ت) في اختبار معنوية β_1 = صفر الذي أشرنا إليه سابقا .

(٢) إذا كــنا نبحــث فـــي مــدى أهمية وجود المتغير المستقل $س_2$ في نموذج الانحدار ، فإننا نقارن بين مجموع المربعات في ظل وجود $س_1$،$س_2$ مع مجموع المربعات فى ظل وجود $س_1$ فقط ، والفرق بين المجموعين يرجع إلى مجموع المـــربعات الإضافي للمتغير $س_2$ أي مساهمة $س_2$ في مجموع المربعات بسبب انحدار ص/$س_1$ ، $س_2$ .. بمعنى آخر فإننا نقارن بين نموذجين ،الأول بدلالة $س_1$ ، $س_2$ والثاني بدلالة $س_1$ فقط أي :

النموذج الأول : $ص = ب + ب_1 س_1 + ب_2 س_2$

النموذج الثاني : $ص = ب + ب_1 س_1$

أما مجموع المربعات بسبب انحدار ص/$س_1$ ، $س_2$ فقد سبق ذكره ، في حين أن مجموع المربعات بسبب انحدار ص/$س_1$ فهو :

$$\text{م.م انحدار ص/س}_1 = \frac{(\text{مجـ ص س}_1 - \text{مجـ ص} \times \text{مجـ س}_1/\text{ن})^2}{\text{مجـ س}^2_1 - (\text{مجـ س}_1)^2/\text{ن}}$$

وبترتيـــب تلـــك المجاميع في جدول تحليل التباين مع م. م. ك ، $خ^2_{ص}$ ، يمكن إيجـــاد مجمـــوع المربعات الإضافي للمتغير $س_2$ والذي يتم اختباره على النحو التالي :

$$\text{ف} = \frac{\text{مجموع المربعات الإضافي أي مساهمة س}_2}{\text{تباين البواقي}}$$

وبالمقارنـــة مـــع ف الجدولية عند درجتي حرية (١ ، ن – ك) ، α يمكن أن نصل إلى قرار بقبول أو رفض الفرض العدمي β_2 = صفر

مثال (٥)

استخدم بيانات مثال (١) والنتائج التي توصلنا إليها في الأمثلة التالية له والمرتبطة به في إجراء :

(١) اختبار β_1 = صفر عند α = ٥% مستخدمًا أسلوب تحليل التباين .

(٢) اختبار β_2 = صفر عند α = ٥% مستخدما أسلوب تحليل التباين .

الحل :

اختبار β_1 = صفر ، اختبار β_2 = صفر ، سبق أن تعرضنا لها عند الحديث عن اختبار معنوية معالم خط الانحدار باستخدام توزيع ت ومثال (٣) كان تطبيقا لذلك . الآن نعيد تلك الاختبارات لكن عن طريق أسلوب تحليل التباين على النحو التالي :

<u>(١) اختبار β_1 = صفر</u>

لتكوين محتوى جدول تحليل التباين ، نجري الآتي :

م. م. ك = مجـ ص٢ − (مجـ ص)٢ / ن = ٢٨ (كما سبق)

م.م. بسبب انحدار ص/س١ ، س٢ = ٢٦,٥ (كما سبق)

$$\text{م.م.بسبب انحدار ص/س}_2 = \frac{(\text{مجـ ص س}_2 - \text{مجـ ص} \times \text{مجـ س}_2 / \text{ن})^2}{\text{مجـ س}_2^2 - (\text{مجـ س}_2)^2 / \text{ن}}$$

$$= \frac{(١٠٩ - ٢٠\times٢٥/٥)^2}{١٢٩ - (٢٥)^2 / ٥} = \frac{٨١}{٤} = ٢٠,٢٥$$

جدول تحليل التباين لاختبار β_1 = صفر

المصدر	م.م	د . ح	م.م.م	ف = نسبة التباين
م.م. انحدار ص/س٢	٢٠,٢٥	١		
الزيادة بسبب س١ أو مساهمة س١	٦,٢٥	١	٦,٢٥	٦,٢٥ / ٠,٧٥ = ٨,٣٣
م.م. انحدار ص/س١ س٢	٢٦,٥	٢		
م.م. د	١,٥	٢	٠,٧٥	
د.د.ك	٢٨	٤		

خطوات الاختبار :

١- الفرض العدمي β_1 = صفر (لا يوجد تأثير للمتغير $س_1$ على التابع ص أي يمكن استبعاده من نموذج الانحدار)

٢- الفرض البديل $\beta_1 \neq$ صفر

٣- ف المحسوبة = ٨,٣٣

٤- ف الجدولية ، ف(٢،٢،٥٪) = ١٨,٥١

٥- المقارنة ، ف المحسوبة أقل من ف الجدولية

٦- القرار : قبول الفرض العدمي بأن β_1 = صفر

أي أن وجود $س_1$ في معادلة خط الانحدار غير مفيد في عملية التنبؤ بالمتغير ص وهذا القرار سبق أن توصلنا إليه في مثال (٣) .

(٢) اختبار β_2 = صفر

م.م.ك = ٢٨ ، م.م. بسبب انحدار ص / $س_1 س_2$ = ٢٦,٥

$$\text{م.م. بسبب انحدار ص / س}_1 = \frac{(\text{مجـ ص س}_1 - \text{مجـ ص} \times \text{مجـ س}_1 / \text{ن})^2}{\text{مجـ س}_1^2 - (\text{مجـ س}_1)^2 / \text{ن}}$$

$$= \frac{(٧٦ - ٢٠ \times ١٥/٥)^2}{٥٥ - (١٥)^2 / ٥} = \frac{(١٦)^2}{١٠} = ٢٥,٦$$

جدول تحليل التباين لاختبار β_2 = صفر

المصدر	م.م	د.ح	م.م.م	نسبة التباين ف
م.م. انحدار ص/$س_1$	٢٥,٦	١		
مساهمة $س_1$	٠,٩	١	٠,٩	$\frac{٠,٩}{٠,٧٥}$ = ١,٢
م.م. انحدار ص/$س_1 س_2$	٢٦,٥	٢		
م.م.د (البواقى)	١,٥	٢	٠,٧٥	
م.م.ك	٢٨	٤		

خطوات الاختبار :

١- الفرض العدمي β_2 = صفر

٢- الفرض البديل β_2 ≠ صفر

٣- ف المحسوبة = ١,٢

٤- ف (٢،١،٥%) = ١٨,٥١

٥- المقارنة : ف المحسوبة أقل من ف الجدولية

٦- القرار : قبول الفرض العدمي

∴ β_2 = صـــفر ، أي أن وجـــود س٢ لا يفيد معنويا في تفسير التغير في ص وبالتالي لا يفيد معنويا في التنبؤ بقيم التابع ص .

مثال (٦) (شامل) :

الجـــدول التالي يعطي كمية الإنتاج من القمح بالطن للفدان (ص) وكمية السماد بالكـــيلو جرام (س١) وكمية المبيدات الحشرية بالرطل (س٢) في إحدى المزارع خلال عشر سنوات متتالية من ١٩٩١ إلى ٢٠٠٠

السنة	١٩٩١	٩٢	٩٣	٩٤	٩٥	٩٦	٩٧	٩٨	٩٩	٢٠٠٠
ص	٤٠	٤٤	٤٦	٤٨	٥٢	٥٨	٦٠	٦٨	٧٤	٨٠
س١	٦	١٠	١٢	١٤	١٦	١٨	٢٢	٢٤	٢٦	٣٢
س٢	٤	٤	٥	٧	٩	١٢	١٤	٢٠	٢١	٢٤

المطلوب :

١- إيجاد معادلة خط الانحدار المتعدد بطريقة المربعات الصغرى .

٢- الخطأ المعياري لتقدير معادلة الانحدار .

٣- معامل التحديد ومعامل التحديد المعدل ومعامل الارتباط المتعدد .

٤- معاملات الارتباط الجزئية ثم حدد أي المتغيرات المستقلة يساهم أكثر في تفسير التغير في التابع ص .

٥- قدر β_1 بفترة ثقة ٩٥%

٦- اختبار β_1 = صفر ، اختبار β_2 = صفر عند α = ٥% مستخدما توزيع ت

٧- اختبار β_1 =صفر ، إختبار β_2 = صفر عند α = ٥% مستخدما توزيع ف (تحليل التباين)

٨- اختبار المعنوية الاجمالية للانحدار المتعدد أي اختبار β_1 = β_2 = صفر مستخدما أسلوب تحليل التباين (α = ٥%)

الحل :

م	ص	س$_1$	س$_2$	ص^2	س$_1^2$	س$_2^2$	ص س$_1$	ص س$_2$	س$_1$ س$_2$
١	٤٠	٦	٤	١٦٠٠	٣٦	١٦	٢٤٠	١٦٠	٢٤
٢	٤٤	١٠	٤	١٩٣٦	١٠٠	١٦	٤٤٠	١٧٦	٤٠
٣	٤٦	١٢	٥	٢١١٦	١٤٤	٢٥	٥٥٢	٢٣٠	٦٠
٤	٤٨	١٤	٧	٢٣٠٤	١٩٦	٤٩	٦٧٢	٣٣٦	٩٨
٥	٥٢	١٦	٩	٢٧٠٤	٢٥٦	٨١	٨٣٢	٤٦٨	١٤٤
٦	٥٨	١٨	١٢	٣٣٦٤	٣٢٤	١٤٤	١٠٤٤	٦٩٦	٢١٦
٧	٦٠	٢٢	١٤	٣٦٠٠	٤٨٤	١٩٦	١٣٢٠	٨٤٠	٣٠٨
٨	٦٨	٢٤	٢٠	٤٦٢٤	٥٧٦	٤٠٠	١٦٣٢	١٣٦٠	٤٨٠
٩	٧٤	٢٦	٢١	٥٤٧٦	٦٧٦	٤٤١	١٩٢٤	١٥٥٤	٥٤٦
١٠	٨٠	٣٢	٢٤	٦٤٠٠	١٠٢٤	٥٧٦	٢٥٦٠	١٩٢٠	٧٦٨
مج	٥٧٠	١٨٠	١٢٠	٣٤١٢٤	٣٨١٦	١٩٤٤	١١٢١٦	٧٧٤٠	٢٦٨٤

(١) معادلة خط الانحدار المتعدد : ص = ب + ب$_1$ س$_1$ + ب$_2$ س$_2$ عن طريق المعادلات الطبيعية الثلاثة التالية ، يمكن إيجاد التقديرات ب$_1$ ، ب$_2$ كمايلي:

مجـ ص = ن ب + ب$_1$ مجـ س$_1$ + ب$_2$ مجـ س$_2$

مجـ ص س$_1$ = ب مجـ س$_1$ + ب$_1$ مجـ س$_1^2$ +ب$_2$ مجـ س$_1$ س$_2$

مجـ ص س$_2$ =ب مجـ س$_2$ + ب$_1$ مجـ س$_1$ س$_2$ +ب$_2$ مجـ س$_2^2$

وبالـــتعويض عـــن المجاميع السابقة بالقيم الموجودة بالجدول ، حيث ن = ١٠ نحصل على ثلاث معادلات في ثلاث مجاهيل ب ، ب$_1$ ،ب$_2$

٥٧٠ = ١٠ ب + ١٨٠ ب$_1$ + ١٢٠ ب$_2$

١١٢١٦ = ١٨٠ ب + ٣٨١٦ ب$_1$ + ٢٦٨٤ ب$_2$

٧٧٤٠ = ١٢٠ ب + ٢٦٨٤ ب$_1$ + ١٩٤٤ ب$_2$

وبحل تلك المعادلات بأي طريقة جبرية كالحذف أو المحددات أو المصفوفات نحصل على القيم التالية : ب = ٣١,٩٨ ، ب$_1$ = ٠,٦٥ ، ب$_2$ = ١,١١

∴ معادلة خط الانحدار المتعدد تصبح على الصورة .

ص = ٣١,٩٨ + ٠,٦٥ س$_1$ + ١,١١ س$_2$

بفحـــص معادلـــة الانحدار التي حصلنا عليها نجد أن تغير س$_1$ (كمية السـماد) بوحـــدة واحدة يؤدي إلى تغير ص (كمية الإنتاج من القمح) بمقدار ٠,٦٥ وحـــدة ، وذلك بغض النظر عن قيمة س$_2$ (كمية المبيدات الحشرية) ، وينطبق نفس المفهوم بالنسبة لمعامل الانحدار الجزئي ب$_2$ = ١,١١ ، بمعنى أن تغـــير كمـــية المبيدات الحشرية بوحدة واحدة يؤدي إلى تغير كمية الإنتاج من القمح (ص) بمقدار ١,١١ وحدة وذلك بغض النظر عن قيمة س$_1$.

إن إحـــدى المشـــاكل الهامة التي تواجه استخدام الانحدار المتعدد ، هي إمكانـــية وجود ارتباط بين المتغيرات المستقلة ، وتسمى هذه المشكلة بالازدواج الخطي بين المتغيرات المستقلة Multicollinearity ، فإذا كان هذا الارتباط قويـــا ، أدى ذلـــك إلـــى تضاؤل مصداقية معاملات الانحدار الجزئية ، وسوف نتـــناول هذه المشكلة في مرحلة لاحقة عند الحديث عن مشاكل تطبيق الانحدار المتعدد .

(٢) الخطأ المعياري لتقدير معادلة الانحدار

$$\text{خـ} = \sqrt{\frac{\text{مجـ}(\text{ص} - \hat{\text{ص}})^2}{\text{ن} - ٣}} = \sqrt{\frac{\text{مجـ ص}^2 - \text{ب مجـ ص} - \text{ب}_1 \text{ مجـ ص س}_1 - \text{ب}_2 \text{ مجـ ص س}_2}{(\text{ن} - ٣)}}$$

$$= \sqrt{(٣٤١٢٤ - ٣١٫٩٨ \times ٥٧٠ - ٠٫٦٥ \times ١١٢١٦ - ١٫١١ \times ٧٧٤٠) \div (١٠ - ٣)}$$

$$= \sqrt{[\,٣٤١٢٤ - ٣٤١١٠٫٤\,] \div ٧} = \sqrt{١٣٫٦ \div ٧} = ١٫٩٤٣ = ١٫٣٩٤$$

(٣) معامل التحديد ومعامل التحديد المعدل : [معادلة (٨) ، معادلة (١١)]

معامل التحديد $\text{ر}^{٢}$ =

$$\frac{\text{ب}_{١}\,[\text{مجـ ص س}_{١} - \text{مجـ ص} \times \text{مجـ س}_{١}/\text{ن}] + \text{ب}_{٢}\,[\text{مجـ ص س}_{٢} - \text{مجـ ص مجـ س}_{٢}/\text{ن}]}{\text{مجـ ص}^{٢} - (\text{مجـ ص})^{٢}/\text{ن}}$$

$$\text{ر}^{٢} = \frac{٠٫٦٥\,[\,١١٢١٦ - ٥٧٠ \times ١٨٠ / ١٠\,] + ١٫١١\,[\,٧٧٤٠ - ٥٧٠ \times ١٢٠ / ١٠\,]}{٣٤١٢٤ - (٥٧٠)^{٢}/١٠}$$

$$= (٠٫٦٥ \times ٩٥٦ + ١٫١١ \times ٩٠٠) \div ١٦٣٤ = ١٦٢٠٫٤ \div ١٦٣٤$$

$$= ٠٫٩٩١٧$$

أي أن ٩٩٫١٧% من التغيرات الكلية في إنتاج القمح تفسر بسبب وجود كل من السماد ($\text{س}_{١}$) والمبيدات ($\text{س}_{٢}$) أما الباقي ٠٫٨٣ % فترجع إلى عوامل عشوائية غير مفسرة .

معامل الارتباط المتعددة هو الجذر التربيعي لمعامل التحديد .

$$\therefore \; \text{ر}_{(\text{ص}٢١)} = \sqrt{٠٫٩٩١٧} = ٠٫٩٩٥٨$$

معـــامل التحديد المعدل $\bar{\text{ر}}^{٢}$ والذي يأخذ في الاعتبار النقص في درجات الحرية عـــند إضـــافة متغيرات مستقلة لنموذج الانحدار يأخذ شكل المعادلة رقم (١١) السابق ذكرها .

$$\bar{\text{ر}}^{٢} = ١ - (١ - \text{ر}^{٢}) \times \frac{\text{ن} - ١}{\text{ن} - \text{ك}}$$

$$= ١ - (١ - ٠٫٩٩١٧) \times \frac{١٠ - ١}{١٠ - ٣} = ١ - ٠٫٠٠٨٣ \times \frac{٩}{٧}$$

$$= ١ - ٠٫٠١٠٧ = ٠٫٩٨٩٣ = ٩٨٫٩٣\,\%$$

بالطبع قيمة $\bar{\text{ر}}^{٢}$ هي دائما أقل من $\text{ر}^{٢}$

(٤) معاملات الإرتباط الجزئية

يمكـــن حســـاب معـــاملات الارتباط الجزئية بدلالة معاملات الإرتباط الخطية البسيطة على النحو التالي :

$$\text{ر}_{\text{ص}١.٢} = \frac{\text{ر}_{\text{ص}١} - \text{ر}_{\text{ص}٢} \times \text{ر}_{٢١}}{\sqrt{١ - \text{ر}^{٢}_{\text{ص}٢}}\ \sqrt{١ - \text{ر}^{٢}_{٢١}}}$$

$$\text{ر}_{\text{ص}٢.١} = \frac{\text{ر}_{\text{ص}٢} - \text{ر}_{\text{ص}١} \times \text{ر}_{٢١}}{\sqrt{١ - \text{ر}^{٢}_{\text{ص}١}}\ \sqrt{١ - \text{ر}^{٢}_{٢١}}}$$

حيـــث $\text{ر}_{\text{ص}١}$ ، $\text{ر}_{\text{ص}٢}$ ، $\text{ر}_{٢١}$ هـــي معاملات الارتباط الخطية البسيطة حيث الأرقام ١ تعنى $\text{س}_{١}$ ، ٢ تعني $\text{س}_{٢}$

$$\text{ر}_{\text{ص}١} = \frac{\text{مجـ ص س}_{١} - \text{مجـ ص} \times \text{مجـ س}_{١} / \text{ن}}{\sqrt{[\text{مجـ ص}^{٢} - (\text{مجـ ص})^{٢}/\text{ن}]\ [\text{مجـ س}^{٢}_{١} - (\text{مجـ س}_{١})^{٢}/\text{ن}]}}$$

$$= \frac{١١٢١٦ - ٥٧٠ \times ١٨٠ / ١٠}{\sqrt{[٣٤١٢٤ - (٥٧٠)^{٢}/١٠]\ [٣٨١٦ - (١٨٠)^{٢}/١٠]}}$$

$$= \frac{٩٥٦}{\sqrt{١٦٣٤ \times ٥٧٦}} = \frac{٩٥٦}{٩٧٠,١} = ٠,٩٨٥٤$$

$\text{ر}_{\text{ص}٢}$: بوضع $\text{س}_{٢}$ مكان $\text{س}_{١}$ في العلاقة السابقة والتعويض نجد أن

$$\text{ر}_{\text{ص}٢} = \frac{٧٧٤٠ - ٥٧٠ \times ١٢٠ / ١٠}{\sqrt{١٦٣٤ \times [١٩٤٤ - (١٢٠)^{٢}/١٠]}}$$

$$= \frac{٩٠٠}{\sqrt{١٦٣٤ \times ٥٠٤}} = \frac{٩٠٠}{٩٠٧,٤٩} = ٠,٩٩١٧$$

$$\text{ر}_{٢١} = \frac{\text{مجـ س}_{١}\text{س}_{٢} - \text{مجـ س}_{١} \times \text{مجـ س}_{٢} / \text{ن}}{\sqrt{(\text{مجـ س}^{٢}_{١} - (\text{مجـ س}_{١})^{٢}/\text{ن})\ (\text{مجـ س}^{٢}_{٢} - (\text{مجـ س}_{٢})^{٢}/\text{ن})}}$$

وبالتعويض مع الاستعانة بالنتائج فى العلاقات السابقة :

$$ر_{٢١} = \frac{٢٦٨٤ - ١٨٠ \times ١٢٠/١٠}{\sqrt{٥٠٤ \times ٥٧٦}} = \frac{٥٢٤}{٥٣٨,٨} = ٠,٩٧٢٥$$

وبالتعويض عن $ر_{ص١}$ ، $ر_{ص٢}$ ، $ر_{٢١}$ فى معاملات الارتباط الجزئية نصل إلى :

$$ر_{ص١.٢} = \frac{٠,٩٨٥٤ - ٠,٩٩١٧ \times ٠,٩٧٢٥}{\sqrt{١ - ٠,٩٨٣٥}\ \sqrt{١ - ٠,٩٤٥٨}} = ٠,٧٠٢٣ = ٧٠,٢٣\%$$

$$ر_{ص٢.١} = \frac{٠,٩٩١٧ - ٠,٩٨٥٤ \times ٠,٩٧٢٥}{\sqrt{١ - ٠,٩٧١}\ \sqrt{١ - ٠,٩٤٥٨}} = ٠,٨٤٣٤ = ٨٤,٣٤\%$$

وحيــث أن $ر_{ص٢.١}$ أكبر من $ر_{ص١.٢}$ ، فهذا يعنى أن المتغير $س_٢$ (كمية المبيدات الحشرية) أكثر أهمية من $س_١$ (كمية السماد) فى تفسير التغيرات فى إنتاج القمح (ص) .

(٥) تقدير $\beta_١$ بفترة ثقة ٩٥% :

$$\beta_١ = ب_١ \pm ت_{(ن-٣،\alpha/٢)} \times ع(ب_١)$$

حيث : $ب_١ = ٠,٦٥$ ، $ت_{(٠,٠٢٥،٧)} = ٢,٣٦٥$

أمـــا الخطأ المعياري للتقدير (ب١) أى ع (ب١) فهو على شكل المعادلة (٢١) أى :

$$ع^٢_{(ب١)} = خ^٢_{ص} \times \frac{مجـ س^٢_٢ - (مجـ س_٢)^٢/ن}{م} \quad \text{حيث}$$

$$م = (مجـ س^٢_١ - (مجـ س_١)^٢/ن)\ (مجـ س^٢_٢ - (مجـ س_٢)^٢/ن)$$
$$- (مجـ س_١ س_٢ - مجـ س_١ \times مجـ س_٢ /ن)^٢$$
$$= [٣٨١٦ - (١٨٠)^٢/١٠]\ [١٩٤٤ - (١٢٠)^٢/١٠]$$
$$- [٢٦٨٤ - ١٨٠\times١٢٠/١٠]^٢$$

= [٥٧٦ × ٥٠٤] − [٥٢٤]^{٢} = ٢٩٠٣٠٤ − ٢٧٤٥٧٦ = ١٥٧٢٨

أما خ^{٢} ص فقد سبق حسابها وكانت : خ^{٢} ص = ١,٩٤٣

$$\therefore \text{ع}^{٢}_{(\text{ب}_{١})} = ١,٩٤٣ \times \frac{(٥٠٤)}{١٥٧٢٨} = ٠,٠٦$$

∴ الخطأ المعياري ع(ب١) = ٠,٢٤

$\beta_{١}$ = ٠,٦٥ ± ٢,٣٦٥ × ٠,٢٤ = ٠,٦٥ ± ٠,٥٧ = ١,٢٢ ، ٠,٠٨

وبالمثل يمكن إيجاد فترة الثقة للمعلمة $\beta_{٢}$ بعد حساب الخطأ المعياري للتقدير ب٢ وهو على الصورة :

$$\text{ع}^{٢}_{(\text{ب}_{٢})} = \text{خ}^{٢}_{\text{ص}} \times \frac{(\text{مجـ س}^{٢}_{١} - (\text{مجـ س}_{١})^{٢} / \text{ن})}{\text{م}}$$

$$= ١,٩٤٣ \times \frac{٥٧٦}{٥٧٦ \times ٥٠٤ - (٥٢٤)^{٢}} = ٠,٠٧$$

$$\therefore \text{ع}_{(\text{ب}_{٢})} = \sqrt{٠,٠٧} = ٠,٢٦$$

(٦) اختبار $\beta_{١}$ = صفر واختبار $\beta_{٢}$ = صفر باستخدام توزيع ت :

<u>أولاً : اختبار $\beta_{١}$ = صفر</u>

خطوات الاختبار :

١- الفرض العدمى $\beta_{١}$ = صفر

٢- الفرض البديل $\beta_{١} \neq$ صفر

٣- $\text{ت}_{\text{المحسوبة}} = \frac{\text{ب}_{١} - \beta_{١}}{\text{ع}(\text{ب}_{١})} = \frac{\text{ب}_{١}}{\text{ع}(\text{ب}_{١})} = \frac{٠,٦٥}{٠,٢٤} = ٢,٧١$

٤- ت الجدولية : $\text{ت}_{\text{ن}-٣ ، \alpha/٢} = \text{ت}_{٧ ، ٠,٠٢٥} = ٢,٣٦٥$

٥- المقارنة : ت المحسوبة أكبر من ت الجدولية

٦- القرار : رفض الفرض العدمى وقبول الفرض البديل .

∴ $\beta_1 \neq$ صفر (أى أن هناك علاقة حقيقية بين المتغير المستقل $س_1$ والتابع ص ومن المهم وجود هذا المتغير فى نموذج خط الانحدار المتعدد .

ثانياً : اختبار β_2 = صفر

خطوات الاختبار :

١- الفرض العدمى β_2 = صفر

٢- الفرض البديل $\beta_2 \neq$ صفر

٣- ت المحسوبة = $\frac{ب_2 - \beta_2}{ع(ب_2)} = \frac{ب_2}{ع(ب_2)} = \frac{1,11}{0,26} = 4,27$

٤- ت الجدولية : $ت_{0,025 ، 7}$ = ٢,٣٦٥

٥- المقارنة : ت المحسوبة أكبر من ت الجدولية

٦- القرار :رفض الفرض العدمى وقبول الفرض البديل ∴ $\beta_2 \neq$ صفر

(٧) اختبار β_1 = صفر ، اختبار β_2 = صفر باستخدام تحليل التباين :

أولاً : اختبار β_1 = صفر أو اختبار مدى أهمية وجود $س_1$ فى معادلة خط الانحدار

م.م.ك = مجـ $ص^2$ -(مجـ ص)^2 / ن = ٣٤١٢٤ - (٥٧٠)^2 /١٠ = ١٦٣٤

م.م.بسبب انحدار ص/$س_1$ $س_2$= $ب_1$(مجـ ص $س_1$- مجـ ص×مجـ $س_1$ /ن)

+$ب_2$(مجـ ص $س_2$ -مجـ ص×مجـ $س_2$/ن)

= ٠,٦٥ (١١٢١٦ - ٥٧٠ × ١٠/١٨٠)

+ ١,١١ (٧٧٤٠ - ٥٧٠ × ١٠/١٢٠)

= ٠,٦٥ × ٩٥٦ + ١,١١ × ٩٠٠

= ٦٢١,٤ + ٩٩٩ = ١٦٢٠,٤

م.م. بسبب انحدار ص/$س_2$ = $\frac{(مجـ ص س_2 - مجـ ص \times مجـ س_2 /ن)^2}{مجـ س_2^2 - (مجـ س_2)^2 /ن}$

$$= \frac{(٩٠٠)^٢}{١٩٤٤ - (١٢٠)^٢/١٠} = \frac{٨١٠٠٠٠}{٥٠٤} = ١٦٠٧,١٤$$

جدول تحليل التباين

المصدر	م.م	د .ح	م.م.م	نسبة التباين
م.م. انحدار ص/س٢	١٦٠٧،١٤	١		$\frac{١٣,٢٦}{١,٩٤٣} = ٦,٨٢$
مساهمة س٢	١٣,٢٦	١	١٣,٢٦	
م.م. انحدار ص/س١، س٢	١٦٢٠,٤٠	٢		
م.م.د (البواقى)	١٣,٦	٧	١,٩٤٣ = $خ^٢_س$	
م.م.ك	١٦٣٤	٩		

خطوات الاختبار :

١- الفرض العدمى $\beta_١$ = صفر .

٢- الفرض البديل $\beta_١ \neq$ صفر .

٣- ف المحسوبة = ٦,٨٢ .

٤- ف الجدولية : ف (١ ، ٧ ، ٥%) = ٥,٥٩ .

٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .

٦- القرار : رفض الفرض العدمى وقبول الفرض البديل .

∴ $\beta_١ \neq$ صفر أى أنه من المهم بقاء س١ فى معادلة خط الانحدار وهو نفس القرار الذى توصلنا إليه باختبار ت .

ثانياً:اختبار $\beta_٢$ =صفر أو اختبار مدى اهمية وجود س٢ فى معادلة خط الانحدار

م.م.ك = ١٦٣٤

م.م. بسبب انحدار ص/س١، س٢ = ١٦٢٠,٤

$$\text{م.م. بسبب انحدار ص/س}_١ = \frac{(\text{مجـ ص س}_١ - \text{مجـ ص} \times \text{مجـ س}_١/\text{ن})^٢}{\text{مجـ س}^٢_١ - (\text{مجـ س}_١)^٢/\text{ن}}$$

$$= \frac{(١١٢١٦ - ٥٧٠ \times ١٨٠ / ١٠)^٢}{٣٨١٦ - (١٨٠)^٢ / ١٠} = \frac{(٩٥٦)^٢}{٥٧٦} = ١٥٨٦,٧$$

جدول تحليل التباين

المصدر	م.م	د . ح	م.م.م	نسبة التباين
م.م. انحدار ص/س١	١٥٨٦,٧	١		
مساهمة س٢	٣٣,٧	١	٣٣,٧	=١٧,٣٤
م.م. انحدار ص/س١ س٢	١٦٢٠,٤	٢		
م.م.د (البواقى)	١٣,٦	٧	$١,٩٤٣ = غ^٢_{ص}$	
م.م.ك	١٦٣٤	٩		

خطوات الاختبار :

١- الفرض العدمى $\beta_٢$ = صفر .

٢- الفرض البديل $\beta_٢ \neq$ صفر .

٣- ف المحسوبة = ١٧,٣٤ .

٤- ف الجدولية : ف (١ ، ٧ ، ٥%) = ٥,٥٩ .

٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .

٦- القرار : رفض الفرض العدمى وقبول الفرض البديل .

∴ $\beta_٢ \neq$ صفر وهو نفس القرار الذى توصلنا إليه باستخدام توزيع ت .

(٨) اختبار المعنوية الكلية للانحدار : اختبار $\beta_١ = \beta_٢$ = صفر

م.م.ك = مجـ ص^٢ − (مجـ ص)^٢ / ن = ١٦٣٤

م.م. بسبب انحدار ص/س١ س٢ = ١٦٢٠,٤

جدول تحليل التباين

المصدر	م.م	د.ح	م.م.م	نسبة التباين
م.م. انحدار ص/س١ س٢	١٦٢٠,٤	٢	٨١٠,٢	٤١٦,٩٨
م.م.د (البواقي)	١٣,٦	٧	١,٩٤٣ = $خ^2$ س	
م.م.ك	١٦٣٤	٩		

خطوات الاختبار :

١- الفرض العدمي $\beta_1 = \beta_2 =$ صفر .

٢- الفرض البديل $\beta_1 \neq \beta_2 \neq$ صفر .

٣- ف المحسوبة = ٤١٦,٩٨ .

٤- ف الجدولية : ف (٢ ، ٧ ، ٥%) = ٤,٧٤.

٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .

٦- القرار : رفض الفرض العدمي وقبول الفرض البديل .

∴ $\beta_1 \neq \beta_2 \neq$ صفر أي أن هناك تأثير معنوني حقيقي لكل من س١ ، س٢ على المتغير التابع ص .

ملحوظة :

من الممكن حساب قيمة ف بطريقة أخرى عن طريق علاقة ف مع معامل التحديد ، حيث :

$$ف = \frac{ر^2}{1 - ر^2} \times \frac{ن - ك}{ك - 1}$$ حيث ك عدد المعالم المقدرة .

وحيث أن $ر^2$ سبق الحصول عليها (أو يمكن الحصول عليها الآن من جدول تحليل التباين) حيث:

$$ر^2 = \frac{\text{م.م بسبب الانحدار}}{\text{م.م.ك}} = \frac{١٦٢٠,٤}{١٦٣٤} = ٠,٩٩١٧$$

$$\therefore \text{ف} = \frac{\text{٠,٩٩١٧}}{\text{١} - \text{٠,٩٩١٧}} \times \frac{\text{١٠}-\text{٣}}{\text{٣} - \text{١}} = \frac{\text{٠,٩٩١٧}}{\text{٠,٠٠٨٣}} \times \frac{\text{٧}}{\text{٢}} = \text{٤١٨,١٨}$$

وهى تقريباً مساوية للقيمة التى ظهرت فى جدول تحليل التباين .

(١٠) بعض مشاكل استخدام تحليل الانحدار :

عـند تقديـر معـالم نموذج خط الانحدار -سواء البسيط او المتعدد- باسـتخدام طـريقة المربعات الصغرى ، كانت هناك مجموعة من الافتراضات يشـترط تحققهـا كى نستخدم هذه الطريقة . هذه الفروض سبق أن تعرضنا لها ويمكن تلخيصها على النحو التالى :

١- القيمة المتوقعة للحدود العشوائية (خ) = صفر .

٢- تباين الحدود العشوائية مقدار ثابت ويساوى $б^{٢}$.

٣- الحدود العشوائية (خ) تتبع توزيع طبيعي .

٤- الحدود العشوائية غير مرتبطة ببعضها البعض ، بمعنى أن الخطأ العشوائي فى فترة ما غير مرتبط بالخطأ أو بالحد العشوائي فى فترة أخرى .

٥- الحدود العشوائية (خ) والمتغيرات المستقلة أو المفسرة غير مرتبطة ببعضها البعض .

٦- هناك فرض إضافى خاص بالانحدار المتعدد ، وهو عدم وجود ارتباط بين المتغـيرات المسـتقلة أو ما يسمي بالازدواج الخطى ، لأنه لو كان هناك ارتـباط تـام بيـن المتغيرات المستقلة لاستحال تقدير معالم خط الانحدار المتعدد بطريقة المربعات الصغرى .

لاشك أن هذه الفروض قد تتوافر من الناحية العملية أو قد لا تتوافر كلها ، فـإذا مـا توفرت تلك الفروض صلحت طريقة المربعات الصغرى للاستخدام وأعطـت تقديرات ذات دقة إحصائية عالية ، أما إذا لم تتوفر هذه الفروض فإن

طريقة المربعات الصغرى لا تصلح للاستخدام ، ويجب البحث عن طريقة تقدير أخرى . ولكي نتأكد من تحقق تلك الفروض وجب علينا إجراء بعض الاختبارات الإحصائية . من الفروض التي إذا سقطت أو لم تتحقق ترتب عليها مشاكل في التقديرات التي نتحصل عليها ما يلي :

١- مشكلة عدم ثبات تباين الحد العشوائي Heteroscedasticity

٢- مشكلة الازدواج الخطى Multicollinearity

٣- مشكلة الارتباط الذاتي للبواقي Autocorrelation

وتعد هذه المشاكل من أكثر المشاكل ظهوراً في التطبيقات العملية ، لذا يجب على الباحث التحقق من خلو البيانات التي يعالجها عن تلك المشاكل قبل استخدام طريقة المربعات الصغرى في عملية تقدير معالم خط الانحدار البسيط أو المتعدد . ونظراً لما تتطلبه هذه المشاكل من معالجة رياضية متقدمة ، بجانب الإلمام بمستوى متقدم من أساليب التحليل الإحصائية ، فسوف نكتفي بعرض تلك المشاكل في نطاق مختصر ، مكتفين بالتعرف على طبيعة المشكلة والنتائج المترتبة على وجودها وكيفية علاجها ويمكن لمن يرغب في معرفة المزيد الاطلاع على الكتب المتخصصة الوارد ذكرها في قائمة المراجع .

١- مشكلة عدم ثبات التباين Heteroscedasticity

من الفروض التي تقوم عليها طريقة المربعات الصغرى ، ثبات تباين الحد العشوائي ، فإذا سقط هذا الفرض ، أدى ذلك إلى أن تكون التقديرات (ب١ ، ب٢) ذات أخطاء معيارية كبيرة، أي تقديرات متحيزة ، ولأن هذه الأخطاء المعيارية (ع(ب١) ، ع(ب٢)) تدخل في تركيب فترات الثقة وفي اختبارات الفروض الإحصائية فهذا يعني :

- انخفاض معنوية معالم الانحدار β_1 ، β_2 .
- اتساع فترات الثقة للمعالم β_1 ، β_2 وهذا الاتساع يعد عيباً في فترة الثقة للمعالم β_1 ، β_2 .

- ضعف الثقة فى نتائج اختبارات الفروض الخاصة بالمعالم β_1 ، β_2 .
- ضعف الثقة فى القيم المتنبأ بها للمتغير التابع بفترة ثقة .

ويمكـــن الكشف عن تحقق أو عدم تحقق فرض ثبات التباين للحدود العشوائية ، باتباع الخطوات التالية :

(اختبار جولد فيلد – كوانت) (Goldfeld & Quandt).

١– ترتـــب البـــيانات ترتيـــباً تصاعدياً (من الأصغر إلى الأكبر) وفق المتغير المستقل س .

٢– تحـــذف بعـــض المشاهدات الوسيطية (خمس مشاهدات مثلا) بحيث يتبقي مجموعتيـــن متساويتين من المشاهدات ، تحسب لكل مجموعة معادلة خط انحدار ، أى معادلة خط انحدار لقيم س الصغرى ومعادلة خط انحدار لقيم س الكـــبرى (إذا لم يتم حذف أية مشاهدات وسيطية يظل الاختبار صحيحاً لكن قوته فى الكشف عن اختلاف التباين تكون أقل) .

٣– مـــن معادلة خط الانحدار الأولى ، تحسب متوسط مجموع مربعات البواقى $خ^2_1$ ،ومـــن معادلة خط الانحدار الثانية ، تحسب متوسط مجموع مربعات البواقى $خ^2_2$.

٤– أوجد النسبة ف حيث ف = $خ^2_2 \div خ^2_1$ ...(٢٨)

والناتج هو متغير عشوائي يتبع توزيع ف بدرجات حرية (ن–و – ٢ك) ÷ ٢ حيث : ن : العدد الإجمالى للمشاهدات فى التجربة كلها ، و: عدد المشاهدات المحذوفة ، ك: عدد المعالم المقدرة.

٥– تقـــارن النســـبة ف مع قيمة جدولية مستخرجة من جدول توزيع ف لنفس درجات الحرية السابقة .

٦– القـــرار : أ– إذا كانت ف المحسوبة أقل من ف الجدولية يقبل الفرض العدمـــى بأنـــه لا توجد مشكلة عدم ثبات التباين أى أن فرض ثبات التباين متحقق فى البيانات .

ب- إذا كانت ف المحسوبة أكبر من ف الجدولية (أو تساويها)يرفض الفرض العدمى ويقبل الفرض البديل بأن هناك مشكلة عدم ثبات التباين أى أن فرض ثبات التباين غير متحقق فى البيانات .

مثال (٧) :

بفرض أن العلاقة الانحدارية بين الأجور (ص) وعدد العاملين (س) فى ٣٠ شركة كانت على الصورة التالية : ص = ١٧,٥ + ٠,٠٩ س . وضح كيف يتم اختبار وجود مشكلة عدم ثبات التباين على فرض توفر بيانات كاملة عن س ، ص .

الحل :

يمكن اختبار وجود مشكلة عدم ثبات التباين ، بترتيب البيانات الأصلية وفق المتغير التفسيري س ترتيباً تصاعدياً ، وإجراء انحدارين منفصلين : الأول لقيم س الصغيرة والثانى لقيم س الكبيرة ، ثم نختبر نسبة مجموع مربعات الخطأ للانحدار الثانى إلى مجموع مربعات الخطأ للانحدار الأول (أى $خ^٢_٢ ÷ خ^٢_١$) ويستخدم توزيع ف عند درجات حرية (ن – و – ٢ك) ÷٢ وهذا الاختبار ويسمي باختبار جولد فيلد – كوانت (Goldfeld & Quandt) يعد مناسباً تماماً للعينات الكبيرة (ن >٣٠) .

بإهمال المشاهدات الست الوسطى مثلاً وتقسيم الباقى إلى مجموعتين ، كل مجموعة من ١٢ مشاهدة وبحساب معادلة خط الانحدار للمجموعة الأولى ، ومعادلة خط الانحدار للمجموعة الثانية ، لنفرض أننا توصلنا إلى النتائج التالية:

للمجموعة الأولى : ص = ٩ + ٠,٦ س ، $خ^٢_١$ = ٠,٥١

للمجموعة الثانية : ص = ١٢ + ٠,٢ س ، $خ^٢_٢$ = ٣,٢

وبحساب النسبة ف حيث ف = $خ^٢_٢ ÷ خ^٢_١$ = ٣,٢ ÷ ٠,٥١ = ٦,٢٧

وحيث أن ف الجدولية عند درجات الحرية = (ن –و–٢ك)÷ ٢ = (٣٠ –٦–٤) ÷٢ = ١٠ هى ف (١٠ ، ١٠ ، ٥%) = ٢,٩٧ .

وحيـــث أن ف المحســـوبة (٦,٢٧) تتجاوز ف الجدولية ، فإننا نرفض الفرض العدمى ونقبل البديل ، أى نقبل بوجود مشكلة عدم ثبات التباين . بالطبع هـــناك عدة طرق لمعالجة مشكلة عدم ثبات التباين لكن هذا الأمر خارج نطاق العرض الحالي .

٣– مشكلة الازدواج الخطى (مشكلة الارتباط بين المتغيرات المستقلة) Multicollinearity

الارتـــباط بيـــن المتغيرات المستقلة هو أحد المشاكل التى تظهر نتيجة لاختلال أحد فروض طريقة المربعات الصغرى ، وهذه المشكلة بالطبع لا توجد فـــى حالة الانحدار البسيط لأنه يشمل متغير مستقل واحد ، ولكن توجد فقط فى حالة الانحدار المتعدد أىَ الذى يشمل عدة متغيرات مستقلة ، وللتوضيح إذا كان نموذج الانحدار المتعدد يشمل متغيرين مستقلين على الصورة : ص = ب + ب $_١$ س$_١$ + ب$_٢$ س$_٢$ فإن مشكلة الازدواج الخطى تصل إلى حدها الأقصي إذا كان هناك ارتباط تام بين المتغيرات المستقلة س$_١$ ، س$_٢$ قيمته $\pm$ ١ ، وتنعدم مشكلة الازدواج الخطى إذا كان الارتباط بين المتغيرات المستقلة مساوياً للصفر ، وهنا تســـمي المتغـــيرات المستقلة (أو التفسيرية) بالمتغيرات المتعامدة . من الناحية العملـــية نـــادراً ما يتحقق أحد الاحتمالين السابقين ، ولكن ما يحدث هو وجود درجـــة من الارتباط الخطى بين المتغيرات التفسيرية أكبر من الصفر وأقل من الواحد الصحيح.

وهـــناك العديد من الأسباب لظهور مشكلة الازدواج الخطى ، لا مجال لذكرها الآن ويمكن الرجوع إلى الكتب المتخصصة للإلمام بها .لكن ما يهمنا هو النـــتائج التـــي قد تترتب على وجود ازدواج خطى ، أي ارتباط بين المتغيرات التفسيرية تتراوح قيمته بين الصفر والواحد .

يترتب على وجود ازدواج خطى كبر حجم الأخطاء المعيارية للتقديرات ($ب_١$ ، $ب_٢$) أى تصبح تقديرات متحيزة ، ومن شأن ذلك أن يؤثر على فترات الثقة وعلى اختبارات الفروض الخاصة بالمعالم $\beta_١$ ، $\beta_٢$ ، إذ تتسع فترات الثقة لتقديرات المعالم ، وهذا يعد عيبا او قصوراً فى فترات الثقة ، كذلك تقل درجة الـثقة فى القرارات التى تتخذ بشأن معنوية المعالم $\beta_١$ ، $\beta_٢$. وهناك العديد من الأساليب الإحصائية التى تستخدم للكشف عن وجود الازدواج الخطى منها:

١- اختبار كلاين — Klein Test

٢- اختبار الارتباط الجزئي — Partial Correlation Test

٣- اختبار فارار-جلوبر — Farrar – Glabuer Test

ويمكن مراجعة الكتب المتخصصة للإلمام بها .

٣- مشكلة الارتباط الذاتى للبواقى Autocorrelation

يقصـد بالارتباط الذاتى وجود ارتباط بين الحدود العشوائية المتتالية $خ_و$ عـبر فـترات زمنية متتالية ، ووجود هذا الارتباط الذاتي يخل بأحد الفروض الأساسية التى تقوم عليها طريقة المربعات الصغرى وهذا الارتباط الذاتى غالباً ما يظهر عند تحليل بيانات السلاسل الزمنية Times series . وقياس الارتباط الذاتـى لا يخـتلف عن قياس الارتباط العادى والذي يتم بين متغيرين ، إلا أن الارتـباط الذاتـى يقيس الارتباط بين القيم المتتالية لنفس المتغير $خ_و$ وهو على الصورة :

$$ر = \frac{مجـ\ خ_و\ خ_{و-١}}{\sqrt{مجـ\ خ^٢_و\ \ مجـ\ خ^٢_{و-١}}} \quad ...(٢٩)$$

حيث $خ_و$: الحد العشوائي فى الفترة الزمنية و .

$خ_{و-١}$: الحد العشوائي فى الفترة الزمنية التى تسبق و مباشرة .

ر : معامل الارتباط الذاتى .

ومبدئــياً يمكــن التعرف على وجود ارتباط ذاتي بين حدود البواقى ، برصد قيم $خ_و$ مع الزمن بيانياً ، فإذا أخذت هذه البواقى شكلاً منتظماً كالدورات ، لكــان ذلــك دليلاً على وجود مشكلة ارتباط ذاتى بين حدود البواقى أى حدود الخطأ العشوائي .

وهناك العديد من الأسباب التي تؤدي إلى حدوث الإرتباط الذاتي منها : اهمــال بعــض المتغــيرات المســتقلة وعدم ادراجها ضمن نموذج الانحدار ، افــتراض صــيغة رياضــية خاطئة للعلاقة بين المتغيرات ، فإذا كانت العلاقة الحقيقية بين المتغيرات هي علاقة منحنى ، واستخدم الباحث علاقة خطية ، فمن شأن هذا ظهور الارتباط الذاتي .. الخ غير ذلك من الأسباب . لكن ما يهمنا هو كيفــية الــتعرف علـــى وجـــود أو عدم وجود إرتباط ذاتي . هناك العديد من الاختبارات الإحصائية يمكن الاستعانة بها للكشف عن هذه المشكلة منها :

١- اختبار فون – نيومان Von Neumann Test

٢- اختبار ديربن – واطسون Durbin – Watson Test

والاختــبار الثاني من أهم وأكثر الاختبارات شيوعاً بين الاحصائين ، ويشترط لاستخدامه أن يكون حجم العينة ١٥ أو أكثر .

<u>اختبار ديربن – واطسون</u>

يتم استخدام هذا الاختبار وفق الخطوات التالية :

١- الفرض العدمي : ρ = صفر   لا يوجـــد ارتباط ذاتي بين البواقي في مجتمع الدراسة ، ρ : معامل الارتباط الذاتي في المجتمع .

٢- الفرض البديل : $\rho \neq$ صفر

٣- احصاء ديربن – واطسون (د) على الصورة

$$د = \frac{مجـ \left(خ_و - خ_{و-١} \right)^٢}{مجـ ح_و^٢} \quad ...(٣٠)$$

٤- القيمة الجدولية للإحصاء (د) تستخرج من جدول ديربن واطسون(موجود في نهاية الكتاب) إما عند مستوى معنوية ١% أو عند مستوى معنوية ٥% والقيمة الجدولية هنا عبارة عن حدين : حد أدنى وحد أعلى أي د١ ، د٢ (d_u ، d_l) يتم الوصول إليهما بدلالة حجم العينة (ن) وعدد المتغيرات التفسيرية فقط . وقيمة إحصاء ديربن – واطسون (د) تتراوح بين صفر ، ٤ . فإذا كانت د قريبة من (٢) لا يكون هناك ارتباط ذاتي . والشكل التالي يوضح قيم د التي عندها يوجد ارتباط ذاتي موجب أو سالب والمناطق التي لا يمكن اتخاذ قرار عندها .

ارتباط ذاتي سالب	قرار غير حاسم	لا يوجد ارتباط ذاتي	قرار غير حاسم	ارتباط ذاتي موجب

٤ ٤-د١ ٤-د٢ ٢ د٢ د١ صفر

٥- المقارنة والقرار : بمقارنة قيمة وسيلة الاختبار (د) مع القيم الجدولية وعند مستوى المعنوية المحدد ، يمكن اتخاذ القرار بقبول أو رفض الفرض العدمي .

بالطبع وكما ذكرنا من قبل ، فإن وجود ارتباط ذاتي بين حدود البواقي من شأنه أن تصبح الأخطاء المعيارية لتقديرات المربعات الصغرى : ع (ب١) ، ع (ب٢) كبيرة أي متحيزة . وحيث أن تلك الأخطاء المعيارية تدخل في تركيب فترات الثقة وفي اختبارات الفروض للمعالم β_1 ، β_2 فإن ذلك من شأنه أن تصبح فترات الثقة غير دقيقة نظراً لإتساع حديها ، كما أن درجة الثقة في نتائج اختبارات الفروض لا يعول عليها بدرجة كبيرة .

مثال (٨)

بفـــرض أنك حصلت على المبيعات التالية عن احدى الظواهر خلال الفترة من ١٩٨٥ إلى ٢٠٠٠ ، وبفرض أن هناك متغير مستقل واحد فقط .

والمطلـــوب دراسة ما إذا كان هناك ارتباط ذاتي بين البواقي أم لا عند مستوى معنوية ٥%

الحل :

السنة	القيمة الفعلية ص	القيمة التقديرية ص^	خ و =ص−ص^	خ²و	خ و−خ و-١	(خ و−خ و-١)²
١٩٨٥	٨١	٨٢	-١	١	----	----
٨٦	٨٢	٨٥	-٣	٩	-٢	٤
٨٧	٩٥	٩٢	+٣	٩	٦	٣٦
٨٨	٧٤	٧٤	٠	٠	-٣	٩
٨٩	٨٣	٨٢	+١	١	١	١
٩٠	٩٥	٩٧	-٢	٤	-٣	٩
٩١	٩٧	٩٥	+٢	٤	٤	١٦
٩٢	١٠٢	١٠٥	-٣	٩	-٥	٢٥
٩٣	١٠٦	١٠٨	-٢	٤	١	١
٩٤	٩٩	١٠١	-٢	٤	٠	٠
٩٥	١٠٠	٩٨	+٢	٤	٤	١٦
٩٦	١٠٢	١٠١	+١	١	-١	١
٩٧	١٠٨	١٠٦	+٢	٤	١	١
٩٨	١١٠	١١٢	-٢	٤	-٤	١٦
٩٩	١١٥	١١٠	+٥	٢٥	٧	٤٩
٢٠٠٠	١١٥	١١٦	-١	١	-٦	٣٦
				٨٤		٢٢٠

لدراســـة الارتـــباط الذاتـــي للبواقي ، علينا بحساب عمود البواقي خ و ، حيث خ و = ص − ص^ = القـــيمة الفعلية − القيمة التقديرية ، ثم حساب الفرق بين

$خ_و$ ، $خ_{و-1}$ أي الفرق بين كل قيمة والقيمة السابقة لها مباشرة وهذا ما توضحه الأعمدة في الجدول السابق

وعن طريق إحصاء ديربن – واطسون الموضح بالمعادلة المعادلة (٣٠) :

$$د = \frac{مجـ (خ_و - خ_{و-1})^2}{مجـ خ_و^2} = \frac{٢٢٠}{٨٤} = ٢,٦٢$$

خطوات الاختبار :

١- الفرض العدمي : ρ = صفر

٢- الفرض البديل : ρ ≠ صفر

٣- د (إحصاء ديربن – واطسون) = ٢,٦٢

٤- القيمة الجدولية عند α = ٥% ، حجم العينة ن = ١٦ ، وعدد المتغيرات التفسيرية =١ نجد أن $د_1$ = ١,١ ، $د_2$ = ١,٣٧ وبتوقيع تلك القيم على الخط الذي يوضح حدود إحصاء ديربن واطسون

ارتباط ذاتي سالب	قرار غير حاسم	لا يوجد ارتباط ذاتي	قرار غير حاسم	ارتباط ذاتي موجب

٤ ٢,٩ ٢,٦٣ ٢ $د_2$ = ١,٣٧ $د_1$ = ١,١ صفر

٦- المقارنة : بتوقيع قيمة د = ٢,٦٢ على الشكل السابق نجدها تقع في المنطقة : لا يوجد ارتباط ذاتي بين حدود البواقي . (لا حظ أن ٢,٦٢ قريبة جدا من ٢,٦٣ وهي منطقة قرار غير حاسم ، لذا يفضل في هذه الحالة تكبير حجم العينة) .

مثال (٩):

(أ) بفرض أنه في نموذج الانحدار البسيط ولعينة من ٢٠ مشاهدة ، كانت معادلة خط الانحدار على الصورة التالية : ص = -٥٦ + ٠,١٣ س وكانت قيمة إحصاء ديربن واطسون هي: د = ٠,٦٥ هل هناك دليل على وجود ارتباط ذاتي بين البواقي؟ (α=٥%)

(ب) بفرض أن معادلة خط الانحدار المتعدد لعينة من ٢٠ مشاهدة على الصورة التالـية . ص = ٧,١٩ + ٠,٢٤ $س_١$ − ٠,٩٩ $س_٢$ وكانت قيمة احصاء ديربن – واطسون هي د=١,٥٤ ، فهل هناك دليل على وجود إرتباط ذاتي بين البواقي ؟ (α = ٥%)

الحل :

(أ) حيــث أن د = ٠,٦٥ أقل من $د_١$ = ١,٢ عند مستوى معنوية ٥% مع ن = ٢٠ ، ك = ١ ، فهناك دليل على وجود ارتباط ذاتي موجب .

(ب)حيث أن د = ١,٥٤ أكبر من $د_٢$ = ١,٥٣ عند مستوى معنوية ٥% مع عينة ن = ٢٠ ، ك = ٢ ، فليس هناك دليل على وجود ارتباط ذاتي .

(لاحظ أن القيم $د_١$ ، $د_٢$ تم الحصول عليها من جدول ديربن –واطسون الموضح فــي نهايــة الكــتاب وذلك عند حجم العينة ن ، وعدد المتغيرات التفسيرية ك ومستوى المعنوية α).

مثال (١٠) نمط محلول

بفــرض أن (ص) تمــثل درجات مستوى الأداء ، $س_١$: المعدل التراكمي في مرحلة البكالوريوس ، $س_٢$: عدد سنوات الخبرة في العمل لعينة من ٢٠ عامل.

الحل :

١− معادلة خط الانحدار المتعدد :

بالتعويض في المعادلات الطبيعية الثلاث ، نصل إلى :

١٣١ = ٢٠ ب + ١٤٧ $ب_١$ + ١١٠ $ب_٢$

١٠١٢ = ١٤٧ ب + ١٢٦١ $ب_١$ + ٨٧٣ $ب_٢$

٧٩٧ = ١١٠ب + ٨٧٣ $ب_٢$ + ٧٨٨ $ب_٢$

وبحل تلك المعادلات بطريقة الحذف أو المحددات أو المصفوفات ، نصل إلى :

ب = ٣,٤٩ ، $ب_١$ = ٠,١٤ ، $ب_٢$ = ٠,٣٦٨٥ = ٠,٣٧

.:. معادلة خط الانحدار هي : ص = ٣,٤٩ + ٠,١٤ $س_١$ + ٠,٣٧ $س_٢$

م	ص	س١	س٢	ص²	س١²	س٢²	ص س١	ص س٢	س١ س٢
١	٥	٣	١	٢٥	٩	١	١٥	٥	٣
٢	٤	٢	٣	١٦	٤	٩	٨	١٢	٦
٣	٤	٤	٢	١٦	١٦	٤	١٦	٨	٨
٤	٩	١٢	٨						
٥	٨	١١	٧						
٦	٩	٨	٤						
٧	٧	٩	١٠						
٨	٨	٧	٥						
٩	٥	٦	٦						
١٠	٦	٥	٣						
١١	٨	٤	٩						
١٢	٤	٨	٤						
١٣	٧	٣	٧						
١٤	٦	١٢	٦						
١٥	٨	٩	٨						
١٦	٥	٨	١						
١٧	١٠	١١	١١						
١٨	٧	٧	٩						
١٩	٦	٨	٥	٣٦	٦٤	٢٥	٤٨	٣٠	٤٠
٢٠	٥	١٠	١	٢٥	١٠٠	١	٥٠	٥	١٠
	١٣١	١٤٧	١١٠	٩٢١	١٢٦١	٧٨٨	١٠١٢	٧٩٧	٨٧٣

(٢) الخطأ المعياري للتقدير خ ص = $\sqrt{\text{مجـ } (\text{ص} - \text{ص}^{\wedge})^{2} \div (\text{ن}-3)}$

$$\text{أو } خ_{ص} = \sqrt{\frac{(\text{مجـ ص}^{٢} - \text{ب مجـ ص} - ب_{١}\,\text{مجـ }س_{١}\,ص - ب_{٢}\,\text{مجـ }س_{٢}\,ص)}{(ن-٣)}}$$

$$= \sqrt{\frac{٢٧,٧٦٤}{١٧}} = \sqrt{١,٦٣٣٢} = ١,٢٨$$

(٣) معـــامل الـــتحديد المتعدد ($ر^{٢}$) وباستخدام المعادلة (٩) نصل إلى : $ر^{٢}$ = ٠,٥٥٧١ أما معامل الإرتباط المتعدد= ٠,٧٤٦ في حين أن معامل التحديد المعدل $\bar{ر}^{٢}$ فهو :

$$\bar{ر}^{٢} = ١ - (١ - ر^{٢}) \times \frac{ن-١}{ن-ك}$$

$$= ١ - (١ - ٠,٥٥٧١) \times \frac{١٩}{١٧} = ١ - ٠,٤٩٥ = ٠,٥٠٥$$

(٤) معاملات الإرتباط الجزئية :

معاملات الارتباط الجزئية $ر_{ص١.٢}$ ، $ر_{ص٢.١}$ وفق المعادلات (١٨) ، (١٩) ، تقتضي معرفة معاملات الإرتباط البسيطة $ر_{ص١}$ ، $ر_{ص٢}$ ، $ر_{٢١}$ وهي على النحو التالي :

$ر_{ص١}$ = ٠,٤٦١ ، $ر_{ص٢}$ = ٠,٧١٣ ، $ر_{٢١}$ = ٠,٣٥٥ بالـــتعويض في (١٨) ، (١٩)

$ر_{ص١.٢}$ = ٠,٣١٥٩ ، $ر_{ص٢.١}$ = ٠,٦٦١٥

(٥) الخطأ المعياري لمعاملات الانحدار الجزئية : $ع(ب_{١})$ ، $ع(ب_{٢})$:-

$$ع^{٢}(ب_{١}) = خ^{٢}_{ص} \times \frac{(\text{مجـ }س^{٢}_{٢} - (\text{مجـ }س_{٢})^{٢}/ن)}{[\text{مجـ }س^{٢}_{١} - (\text{مجـ }س_{١})^{٢}/ن]\,[\text{مجـ }س^{٢}_{٢} - (\text{مجـ }س_{٢})^{٢}/ن] - [\text{مجـ }س_{١}س_{٢} - \text{مجـ }س_{١} \times \text{مجـ }س_{٢}/ن]^{٢}}$$

$$= ١,٦٣٣ \times \frac{١٨٣}{١٨٣ \times ١٨٢,٥٥ - ٤١٦٠,٢٥} = ١,٦٣٣ \times \frac{١٨٣}{٢٩٢٤٦,٤} = ٠,٠١٠٢$$

∴ الخطأ المعياري $ع(ب_{١}) = ٠,١٠١$

بالمثل ع^{٢} (ب$_{٢}$) = ١,٦٣٣ × $\frac{١٨٢,٥٥}{٢٩٢٤٦,٤}$ = ٠,٠١٠٢

∴ الخطأ المعياري ع (ب$_{٢}$) = ٠,١٠١

(٦) احسب فترة الثقة ٩٥% لمعاملات الانحدار الجزئية :

$\beta_{١}$ = ٠,١٤ ± ٢,١١ × ٠,١٠١

$\beta_{٢}$ = ٠,٣٦٨٥ ± ٢,١١ × ٠,١٠١

(٧) اختبر فرض العدم القائل بأن معامل انحدار المجتمع $\beta_{١}$ يساوي الصفر عند α = ٠,٠٥ ثم نفذ الاختبار على المعامل $\beta_{٢}$ باستخدام توزيع ت

ت$_{١}$ = $\frac{ب_{١} - \beta_{١}}{ع\ (ب_{١})}$ = $\frac{٠,١٤ - صفر}{٠,١٠١}$ = ١,٣٨٦

ت$_{٢}$ = $\frac{ب_{٢} - \beta_{٢}}{ع\ (ب_{٢})}$ = $\frac{٠,٣٦٨٥ - صفر}{٠,١٠١}$ = ٣,٦٤٩

يلاحظ أن ت$_{١}$ أقل من ت الجدولية (٢,١١) ومن ثم يقبل الفرض العدمي بينما ت$_{٢}$ أكبر من ت الجدولية ومن ثم يقبل الفرض البديل .

(٨) اختبار العلاقة الإجمالية باستخدام تحليل التباين

م.م.ك = ٦٢,٩٥

م.م.بسبب انحدار ص/س$_{١}$ ، س$_{٢}$=٠,١٤×٤٩,١٥+٠,٣٧×٧٦,٥= ٣٥,١٨٦

جدول تحليل التباين

المصدر	م.م.	د.ح	م.م.م	نسبة التباين
انحدار ص/س$_{١}$س$_{٢}$	٣٥,١٨٦	٢	١٧,٥٩٣	١٠,٧٧
م.م.د	٢٧,٧٦٤	١٧	١,٦٣٣٢	
م.م.ك	٦٢,٩٥	١٩		

وحيث أن ف الجدولية : ف (٢،١٧، ٥%) = ٣,٥٩ يكون القرار قبول الفرض البديل $\beta_{١} \neq \beta_{٢}$ = صفر

(٩) اختبار مدى أهمية وجود $س_١$ في معادلة خط الانحدار

م.م. انحدار ص / $س_٢$ =

$$\frac{[\text{مجـ ص } س_٢ - \text{مجـ ص} \times \text{مجـ } س_٢/ن]^٢}{\text{مجـ } س^٢_٢ - (\text{مجـ } س_٢)^٢/ن} = \frac{(٧٦,٥)^٢}{١٨٣} = ٣١,٩٧٩٥ = ٣٢$$

جدول اختبار معنوية $س_١$

المصدر	م.م	د.ح	م.م.م	نسبة التباين	ف الجدولية
م. انحدار ص/$س_٢$	٣٢	١	٣,١٨٦	١,٩٥١	ف(١ ، ١٧) = ٤,١٥
مساهمة $س_١$	٣,١٨٦	١			
م.م انحدار ص/ $س_١$ $س_٢$	٣٥,١٨٦	٢			
م.م.د	٢٧,٧٦٤	١٧	١,٦٣٣٢		
م.م.ك	٦٢,٩٥	١٩			

وحيث أن ف المحسوبة أقل من ف الجدولية ، يقبل الفرض العدمي أي أن $س_١$ متغير غير هام ولا ينصح ببقائه في معادلة خط الانحدار المتعدد .

(١٠) اختبار مدى أهمية وجود $س_٢$ في معادلة خط الانحدار

$$\text{م.م. انحدار ص}/س_١ = \frac{(٤٩,١٥)^٢}{١٨٠,٥٥} = ١٣,٣٨$$

وبتكوين جدول تحليل التباين ، نجد أن ف المحسوبة = ١٣,٣٥ وهي أقل من ف الجدولـــية (٤,١٥) ممـــا يعني قبول الفرض البديل بأن $\beta_٢ \neq$ صفر أي أنه من الضروري بقاء $س_٢$ في معادلة خط الانحدار .

مثال (١١) : نصف محلول

الجـــدول التالي يبين توزيع الوزن (بالرطل) ص ، الطول (بالبوصة) $س_١$ ، العمر (بالسنوات) $س_٢$ لعينة من ١٢ طالب :

ص	٦٤	٧١	٥٣	٦٧	٥٥	٥٨	٧٧	٥٧	٥٦	٥١	٧٦	٦٨
س$_١$	٥٧	٥٩	٤٩	٦٢	٥١	٥٠	٥٥	٤٨	٥٢	٤٢	٦١	٥٧
س$_٢$	٨	١٠	٦	١١	٨	٧	١٠	٩	١٠	٦	١٢	٩

المطلوب :

١- تقدير معادلة خط الانحدار المتعدد بطريقة المربعات الصغرى

٢- اختبار العلاقة الإجمالية بين المتغيرات الثلاث مستخدما أسلوب تحليل التباين عند مستوى معنوية ٥% .

٣- معامل التحديد ومعامل التحديد المعدل ومعامل الارتباط المتعدد .

٤- اختبر مدى أهمية وجود العمر (س$_٢$) في معادلة خط الانحدار ثم كرر التحليل بالنسبة للطول (س$_١$) عن مستوى معنوية ١% .

الحل :

(١) من بيانات الجدول ، يمكن أن نصل إلى المجاميع التالية .

مجـ ص = ٧٥٣   مجـ س$_١$ = ٦٤٣   مجـ س$_٢$ = ١٠٦

مجـ ص^٢ = ٤٨١٣٩   مجـ س$^٢_١$ = ٣٤٨٤٣   مجـ س$^٢_٢$ = ٩٧٦

مجـ ص س$_١$ = ٤٠٨٣٠   مجـ ص س$_٢$ = ٦٧٩٦   مجـ س$_١$س$_٢$ = ٥٧٧٩

معادلة خط الانحدار المتعدد :

ص = ٣,٦٥١٢ + ٠,٨٥٤٦ س$_١$ + ١,٥٠٦٣ س$_٢$

(٢) م.م.ك = ٤٨١٣٩ − (٧٥٣)٢/١٢ = ٨٨٨,٢٥

م.م.ب = ٠,٨٥٤٦ × ٤٨١,٧٥ + ١,٥٠٦٣ × ١٤٤,٥ = ٦٢٩,٣٦

ف المحسوبة = $\frac{٣١٤,٦٨}{٢٨,٧٦}$ = ١٠,٩٤

ف الجدولية : ف (٢ ، ٩ ، ١%) = ٨,٠٢

(٣) معامل التحديد ر^٢ = $\frac{٦٢٩,٣٦}{٨٨٨,٢٥}$ = ٠,٧٠٨٥ ، ومعامل الارتباط = ٠,٨٤١٧

أما معامل التحديد المعدل $ر^{/٢} = ١ - (١ - ر^٢) \times \frac{ن - ١}{ن - ك}$

$= ١ - ٠,٢٩١٥ \times \frac{١١}{٩} = ٠,٦٤٤$

(٤) اختبار مدى أهمية $س_٢$ (العمر) أي اختبار $\beta_٢$ = صفر

م.م انحدار ص / $س_١ = \frac{(٤٨١,٧٥)^٢}{٣٨٨,٩١٧} = ٥٩٦,٧٤٢$

بعـــد تكوين جدول تحليل التباين ، نجد أن : مساهمة $س_٢$ = ٣٢,٦١٨ ونجـــد أيضاً أن : ف المحسوبة = ١,١٣٤ وهى أقل من ف (١ ، ٩، ٥%) = ٥,١٢ . وعلـــى ذلك تعتبر $س_٢$ غير معنوية ، أى أن وجود العمر على حدة فى معادلة خط الانحدار لا يفيد معنوياً فى تفسير التغير فى الوزن ، وبالتالى لا يفيد معنوياً فى التنبؤ بقيم الوزن فى غيبة الطول .

وعند اختبار مدى أهمية طول القامة ($س_١$) ، أي إختبار $\beta_١$ = صفر نجد أن :

م.م. انحدار ص / $س_٢ = \frac{(١١٤,٥)^٢}{٣٩,٦٦٧} = ٥٢٦,٣٨٨$

وبعـــد تكون جدول تحليل التباين ، نجد أن مساهمة $س_١$ = ١٠٢,٩٧٢ وسنجد أن ف المحسوبة لها = ٣,٥٨ وهي أقل من ف الجدولية (٥,١٢) مما يعنـــي قـــبول الفـــرض العدمي أي أن $س_١$ يعتبر متغير غير معنوي في غيبة المتغـــير $س_٢$ بمعـــنى أنه ليس من المفيد استخدام $س_١$ على حدة في غيبة $س_٢$ لتفسير التغير في الوزن ص . مما تقدم نجد أن التغير في الوزن يفسر في ضوء وجـــود كلا المتغيرين معا $س_١$ ، $س_٢$ وأن أحدهما لا يغني عن الآخر في تفسير التغير في الوزن .

مثال (١٣) : نصف محلول

الجـــدول التالـــي يعطي دخل الفرد الحقيقي بالألف دولار (ص) ، نسبة القـــوة العاملة في الزراعة ($س_١$) ، متوسط سنوات التعليم ($س_٢$) لعدد ١٥ دولة

متقدمة في عام ١٩٨١ . المطلوب :

١- معادلة خط الانحدار المتعدد بطريقة المربعات الصغرى .

٢- الخطأ المعياري لمعادلة خط الانحدار .

٣- تباين التقديرات $ب_1$،$ب_2$ والأخطاء المعيارية لها .

٤- اختبار معنوية المعالم β_1 ، β_2 عند مستوى معنوية ٥% باستخدام توزيع ت وكذلك فترات الثقة لتلك المعالم .

٥- معامل التحديد ومعامل التحديد المعدل .

٦- اختبار المعنوية الإجمالية لمعادلة خط الانحدار باستخدام تحليل التباين .

٧- معاملات الارتباط الجزئية.   والبيانات كما يلي :

ص	٦	٨	٨	٧	٧	١٢	٩	٨	٩	١٠	١٠	١١	٩	١٠	١١
$س_1$	٩	١٠	٨	٧	١٠	٤	٥	٥	٦	٨	٧	٤	٩	٥	٨
$س_2$	٨	١٣	١١	١٠	١٢	١٦	١٠	١٠	١٢	١٤	١٢	١٦	١٤	١٠	١٢

الحل :

١- معادلة خط الانحدار المتعدد :

$$\text{ص} = ٦{,}٢٦ - ٠{,}٣٨\ \text{س}_1 + ٠{,}٤٥\ \text{س}_2$$

٢- تباين البواقي (تباين تقديرات معادلة خط الانحدار) $\text{خ}^2_{\text{ص}}$

$$\text{خ}^2_{\text{ص}} = \frac{\text{مجـ}(\text{ص} - \hat{\text{ص}})^2}{\text{ن} - ٣} = \frac{١٢{,}٢٧٣}{١٢} = ١{,}٠٢$$

∴ الخطأ المعياري $\text{خ}_{\text{ص}}$ = ١,٠٠٩٩ = ١ تقريباً

٣- تباين التقديرات $ب_1$ ، $ب_2$

$ع^2(ب_1) = ٠{,}٠٢$ ، $ع^2(ب_2) = ٠{,}٠١$

أما الأخطاء المعيارية $ع(ب_1) = ٠{,}١٤$ ، $ع(ب_2) = ٠{,}١$

٤- اختبار معنوية المعالم β_1 ، β_2 باستخدام توزيع ت

$$ت_1 = \frac{-٠{,}٣٨}{٠{,}١٤} = ٢{,}٧١ \quad ، \quad ت_2 = \frac{٠{,}٤٥}{٠{,}١} = ٤{,}٥$$

أما ت الجدولية : ت (١٢ ، ٠,٠٢٥) = ٢,١٧٩

∴ β_1 معنوية ، β_2 معنوية

أما فترات الثقة للمعالم β_1 ، β_2 :

β_1 = −٠,٣٨ ± ٢,١٧٩ × ٠,١٤ = −٠,٠٧ ، −٠,٦٩

β_2 = ٠,٤٥ ± ٢,١٧٩ × ٠,١ = ٠,٢٣ ، ٠,٦٧

٥- معامل التحديد ومعامل التحديد المعدل :

ر^2 = ٠,٦٩٣٢ ، ر$^{-2}$ (المعدل) = ٠,٦٤١

٦- اختبار المعنوية الإجمالية لمعادلة خط الانحدار :

$$\text{ف المحسوبة} = \frac{٢٧,٧٢٧ \div ٢}{١٢,٢٧٣ \div ١٢} = ١٣,٥٩$$

وهى أكبر من ف الجدولية(٣,٨٨)عند مستوى معنوية ٥%ودرجات حرية٢،١٢

ويمكن حساب ف بدلالة ر^2 :

$$\text{ف} = \frac{ر^2}{١-ر^2} \times \frac{ن - ك}{ك - ١} = \frac{٠,٦٩٣٢}{١-٠,٦٩٣٢} \times \frac{١٢}{٢} = ١٣,٥٦$$

٧- معاملات الارتباط الجزئية :

هنا يلزم معرفة معاملات الارتباط البسيطة وهى :

ر$_{ص١}$ = −٠,٥٧١٥ ، ر$_{ص٢}$ = ٠,٦٩٨٤ ، ر$_{٢١}$ = −٠,١٨٠١

∴ ر$_{ص١.٢}$ = −٠,٦٣٣١ ، ر$_{ص٢.١}$ = ٠,٨٠٧٢

تمــارين

١- أ- ما هى فروض نموذج الانحدار المتعدد ؟

ب- مــا هــو المقصــود بمعــامل الانحدار الجزئي ومعامل الارتباط الجزئي؟

ج- ما هو المقصود بمعامل التحديد ومعامل التحديد المعدل ؟

د- إذا كانــت قيمة معامل التحديد فى نموذج الانحدار البسيط هى ٠,٧ فهــل تكون قيمة معامل التحديد فى نموذج الانحدار المتعدد أكبر أم أقل من ٠,٧ ؟ ولماذا ؟

هـــ- ما هى أهمية الخطأ المعياري لتقدير معاملات خط الانحدار ؟

٢- ترغب إحدى شركات التأمين على الحياة فى تحديد العلاقة بين حجم مبيعات مندوبيها وبين خبرتهم فى هذا المجال وحجم أسرة كل منهم . سحبت عينة عشــوائية من تسعة مندوبين وسجلت لهم : المبيعات (ص) بالألف جنيه ، الخبرة (س١) بالسنوات ، عدد أفراد الأسرة (س٢) والنتائج هى :

ص	١	٢	٣	٤	٥	٦	٧	٨	٩
س١	٢	١	٣	٣	٤	٥	٦	٥	٧
س٢	١	٢	١	٣	١	٤	٧	٥	٦

<u>المطلوب :</u>

أ- معادلــة خــط انحــدار ص/س١ ، س٢ مستخدماً طريقة المربعات الصغرى .

ب- قدر حجم المبيعات السنوية لمندوب خبرته ١٠ سنوات وأسرته من ٥ أفراد .

ج- حدد الخطأ المعياري للتقدير .

د- احسب معاملات الارتباط الجزئية .

هـ- أوجد فترة الثقة ٩٥% لمعاملات الانحدار β_1 ، β_2 .

و- اختبار فرض العدم القائل بأن كل معامل انحدار جزئي يساوى الصفر فى مقابل أنه لا يساوى الصفر عند مستوى معنوية ٥% .

ز - اختبر العلاقة الإجمالية لمعادلة خط الانحدار مستخدماً أسلوب تحليل التباين .

ح- اختبر مدى أهمية وجود المتغير س١ فى معادلة خط الانحدار .

٣- أجريت دراسة لتحديد العلاقة بين درجات مستوى الأداء (ص) لعشرة مهندسين . وبين الخبرة (س١) وعمر كل واحد منهم (س٢) وكانت كما يلي:

ص	١	٢	٣	٤	٥	٥	٦	٧	٨	٩
س١	١	٣	٢	٥	٥	٤	٧	٦	٩	٨
س٢	٢٤	٢٤	٣٥	٣٢	٣١	٣٨	٣٦	٥٠	٤٢	٤٨

المطلوب :

أ- معادلة خط الانحدار المتعدد ص/س١ ، س٢ .

ب-معامل التحديد ومعامل التحديد المعدل .

ج- اختبار معنوية β_1 عند α = ٥% مستخدماً أسلوب فترة الثقة ثم دعم ذلك مستخدماً توزيع ت .

د - اختبار β_1 = β_2 = صفر باستخدام تحليل التباين ، α = ٥% .

٤- البيانات التالية تمثل عدد مرات الإجازات المرضية (ص) وعدد سنوات الخدمة (س١) وعدد أفراد الأسرة (س٢) لعينة عشوائية من خمس موظفين بإحدى الشركات :

ص	١٠	١٦	١٤	١٥	١٥
س١	١٥	٩	١٣	١١	١٢
س٢	٦	٢	٤	٥	٣

المطلوب :

أ- معادلة خط انحدار ص/$س_1$ ، $س_2$.

ب-عـدد الإجـازات المرضية لموظف بالشركة خدمته ١٥ سنة وعدد فراد أسرته أربعة أفراد.

ج- احسب الخطأ المعياري لتقدير معادلة خط الانحدار .

د- احسب معامل التحديد ومعامل التحديد المعدل ومدلول كل منهما .

هـ- احسب فترة الثقة ٩٥% لمعامل الانحدار β_1 .

و- اختبر معنوية المعامل β_2 عند α = ٥% . مستخدماً توزيع ت ثم دعم ذلك باستخدام أسلوب تحليل التباين .

ز- اختبر معنوية المتغير $س_1$ عند α = ٥% بطريقتين مختلفتين .

٥- أجريت دراسة عن العلاقة بين درجة التحصيل فى مادة الإحصاء (ص) وبين كل من عدد ساعات المذاكرة ($س_1$) والعمر ($س_2$) على عينة من الدارسين وكانت النتائج كما يلي :

الطالب	أ	ب	ج	د	هـ	و
ص	٣٨	٦٧	٤٩	١٢	١٠	٢٤
$س_1$	٩	١١	١٤	٤	٦	٦
$س_2$	٢٠	١٩	٢١	٢٦	٢٤	٣٠

المطلوب :-

أ- معادلة خط أنحدر ص / $س_1$ ، $س_2$

ب- اختبار β_1 = صفر هند α = ٥% مستخدما أسلوب فترة الثقة للعملة β_1 ثم دعم قرارك باستخدام أسلوب اختبار المعنوية (اختبار ت) ثم نـاقش علاقـة هذا الاختبار باختبار مدى أهمية المتغير $س_1$ فى معادلة خط الانحدار .

جـ- معامل التحديد ومعامل التحديد المعدل ثم قارن بينهما .

د- معاملات الارتباط الجزئية .

٦- البيانات التالية تمثل كمية فول الصويا (ص) وكمية المياه المستخدمة فى الرى (س$_{١}$) ، وكمية السماد (س$_{٢}$) لعينة عشوائية من ست أفدنة .

ص$_{ى}$	٢٠	٢٨	٣٢	٢٥	٣٠	٣١
س$_{١ى}$	صفر	٢	٤	١	٣	٥
س$_{٢ى}$	٢	٣	٣	٢	٢	٣

<u>المطلوب :</u>

أ- معادلة خط انحدار ص/س$_{١}$ ، س$_{٢}$

ب- فترة الثقة ٩٥% لكل من معاملي الانحدار الجزئي فى المجتمع .

جـ- اختبار العلاقة الإجمالية لمعادلة خط الانحدار باستخدام أسلوب تحليل التباين .

هـ- اختبار مدى أهمية وجود س$_{١}$ فى معادلة خط الانحدار عند α = ٥% ، ثم كرر التحليل بالنسبة للتغير س$_{٢}$.

٧- بفرض أنك حصلت على البيانات التالية :

ص$_{ى}$	٢٠	٢٨	٤٠	٤٥	٣٧	٥٢	٥٤	٤٣	٦٥	٥٦
س$_{١ى}$	٢	٣	٥	٤	٣	٥	٧	٦	٧	٨
س$_{٢ى}$	٥	٦	٦	٥	٥	٧	٦	٦	٧	٧

<u>المطلوب</u> : تأكد من النتائج التالية :

أ- معادلة خط الانحدار المتعدد : ص=٤,٧٦+٥,٢٩ س$_{١}$+٢,١٣ س$_{٢}$.

ب- الخط المعياري للتقدير = ٧,٠٧١ ، الخطأ المعياري لمعاملات خط الانحدار : ع(ب$_{١}$) = ١,٧٨ ، ع(ب$_{٢}$) = ١٨,٩٥ .

جـ- β_1 معنوية إحصائيا عند α = ٥% .

β_2 ليست معنوية إحصائيا عند α = ٥% .

د- فترة الثقة ٩٥% لمعالم خط الإنحدار هى :

β_1 = ١,٠٨ ، ٩,٥ & β_2 = -٨,١٦ ، ١٢,٤٢

هـ- معامل التحديد ر^2 = ٠,٧٩ ، معامل التحديد المعدل ر$^{-2}$ = ٠,٧٣ .

و- مجموع المربعات بسبب إنحدار ص/س$_1$،س$_2$ = ٦٤٩ .

مجمـوع مـربعات الـبواقى = ٥٠ ، ف المحسوبة (٢،٧) = ١٢,٩٨ ويقبل الفرض البديل β$_1$ ≠ β$_2$ ≠ صفر .

ز- معاملات الارتباط الجزئية : ر$_{ص١.٢}$ = ٠,٧٤ ، ر$_{ص٢.١}$ = ٠,١٨

أى متغـير مسـتقل يسـاهم أكثر فى القدرة التفسيرية للنموذج ؟

الإجابة : س$_1$.

٨- الجـدول التالـي يعطـى بيانات عينة عشوائية من ١٢ أسرة ، تشمل عدد الأطفـال فى الأسرة (ص) ، عدد الأطفال الذين كانوا يرغبون فى إنجابهم وقت الزواج (س$_1$) ، وعدد سنوات التعليم للزوجة (س$_2$) .

الأسرة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
ص	٤	٣	٠	٤	٤	٣	٠	٤	٣	١	٣	١
س$_1$	٣	٣	٠	٢	٢	٣	٠	٣	٢	١	٣	٢
س$_2$	١٢	١٤	١٨	١٠	١٠	١٤	١٨	١٢	١٥	١٦	١٤	١٥

المطلوب :

أ- معادلة خط الإنحدار المتعدد ص / س$_1$، س$_2$

ب- إختبار معنوية β$_1$ ، β$_2$ عند α = ٥% مستخدما توزيع ت.

جـ- أوجد معامل التحديد و معامل التحديد المعدل ثم قارن بينهما .

د- إختبار المعنوية الأجمالية لمعادلة خط الإنحدار عند α = ٥% .

هـ- أوجـد معاملات الأرتباط الجزئية وحدد أى المتغيرات المستقلة يساهم أكثر فى القدرة التفسيرية للنموذج .

ملحوظة : تأكد من الإجابات التالية :

أ- ص = ٦,٩ + ٠,٥٣ س$_1$ – ٠,٣٩ س$_2$.

ب- حيث أن $ت_١$ = ٣,١٢ & $ت_٢$ = -٥,٥٧ ، فإن كلا من $\beta_١$ ، $\beta_٢$ تصبح معنوية إحصائيا عند α = ٥% .

جـ- $ر^٢$ = ٠,٩٢ ، $ر^{-٢}$ = ٠,٩٠ .

د- ف (٢ ، ٩) = ٥١,٣١ .

هـ- $ر_{ص٢.١}$ = ٠,٧١ & $ر_{ص١.٢}$ = ٠,٨٧ بالتالى فإن $س_٢$ تساهم أكثر من $س_١$ فى القدرة التفسيرية للنموذج .

٩- تبين من أحد التقارير أن : (ص) تمثل حجم المبيعات ، ($س_١$) عدد ساعات العمل ، ($س_٢$) عدد أشهر الخبرة وذلك لعينة عشوائية مكونة من ١٠ عمال من أحدى الشركات :

مجـ ص = ١٠٣ مجـ $س_١$ = ١٠٢ مجـ $س_٢$ = ٧٥

مجـ $ص^٢$ = ١٣٦٩ مجـ ص $س_١$ = ٢٨٨ مجـ ص $س_٢$ = ٩٨٩

وبفرض أنك حصلت على معادلة خط الإنحدار المتعدد على الصورة .

ص= -٢,٥ + ١,١٠٣ $س_١$ + ٠,٢٠٧ $س_٢$

<u>المطلوب :</u>

اختبار العلاقة الإجمالية لمعادلة خط الانحدار عند α = ٥% .

١٠- (أ) بغرض أن هناك ٥٠ مشاهدة وأربع متغيرات تفسيرية ، ما هو تعليقك حول الارتباط الذاتى إذا كانت قيمة إحصاء ديرين - واطسون على الصور التالية : (α =٥%)

د= ١,٠٥ ، د= ١,٤ ، د= ٢,٥ ، د= ٣,٩٧

(ب) بفرض أنك حصلت على النموذجين التاليين ، كل منها على عينة من ١٦ مفردة

ص = ٠,٤٥ - ٠,٠٠٤ س ، د= ٠,٨٢٥

ص = ٠,٤٧ - ٠,٠١ $س_١$ + ٠,٠٠٥ $س_٢$ ، د = ١,٢٨

حدد أى النموذجين به ارتباط ذاتى عند α = ٥%

الفصل الخامس
الطـرق اللامعلميــة
Non-Parametric Methods

محتويــــات الفصل:

(١) : مقدمــــة.
(٢) : توزيع كا^{2}.
(٣) : إختبار الإشارة.
(٤) : إختبار مجموع الرتب / إختبار U: مان-هوتينى.
(٥) : إختبار ولكوكسن التى تعتمد على الرتب.
(٦) : تحليل التباين فى اتجاه واحد بالرتب / إختبار كروسكال-والس.
(٧) : معامل إرتباط الرتب لسبيرمان واختبار معنويته.

تمارين

الفصـل الخامس

الطـرق اللامعلميــة

Non-Parametric Methods

(١) مقدمـة :

سبق أن عرفنا الإستنتاج الرياضى بأنه الوصول إلى قرارت بشأن المتغير موضـــوع الدراسة فى المجتمع إعتماداً على نتائج تسجل عن مفردات عشوائية تسحب من ذلك المجتمع ولقد تعرضنا – فى هذا المجال – إلى بعض الإختبارت الإحصـــائية المتعلقة بمتوسط واحد أو الفرق بين متوسطين ثم التى تعتمد على عدة متوسطات وكانت أدوات الإختبارات الإحصائية عبارة عن متغيرات تربط مـــا بيـــن القـــيمة المحسوبة من عينة عشوائية واحدة أو أكثر والقيمة الفرضية للمتوســـط (أو المتوســـطات) – تحـــت شرط صحة الفرض العدمى – بتوزيع احتمالى معين يمكن تعيين إحتمالاته من جداول خاصة بكل توزيع إحتمالى. ولقد كانـــت هذه الإختبارات تفترض – لصلاحية استخدامها – تحقق فروضاً معينة تتصل بنوع القياسات وأنها قياسات كمية لفترة (Interval) أو نسبية (Ratio) وأن توزيع المتغير فى المجتمع يتبع التوزيع المعتاد الطبيعى فضلاً عن شروطاً أخرى تتعلق بتساوى التباينات داخل المعالجات وأن الخطأ العشوائى مستقل فى توزيعه عن المعالجات وأنها تتبع فى توزيعها التوزيع المعتاد.

ومـــع ذلـــك فإننا كثيراً ما نواجه مجالات لا نستطيع فيها أن نجزم بشكل التوزيع وأن أحسن ما يمكن أن نفترض بشأنه أن توزيع مستمر كما أن القياسات قـــد تكـــون نوعـــية (Norminal) أو ترتيبه (Ordinal) فمثلاً لو أن ترتيب مفـــردات عيـــنة عشوائية من الوحدات الصناعية بحسب درجة الأمن الصناعى ومعدل إصابات العمل كانت على النحو التالى (جدول ٥-١):

جدول (١)

التوزيع الترتيبى لعينة عشوائية من الوحدات الصناعية بحسب درجة الأمن الصناعى ومعدل إصابات العمل

الوحدة الصناعية	١	٢	٣	٤	٥	٦	٧	٨
الترتيب بحسب درجة الأمن الصناعى:	٦	٣	٤	١	٧	٢	٨	٥
الترتيب بحسب معدل إصابات العمل:	٥	١	٨	٢	٦	٣	٧	٤

فــإن بــيانات كهــذه لا تعطى تفصيلات كافية عن درجة الأمن الصناعى ومعدل إصابات العمل ولا تعنى بالضرورة أن الوحدة الصناعية رقم (٣) يصل معــدل إصابات العمل بها ثمانية أمثال الوحدة رقم (٢) ولو أن البيانات أعطيت فى صورة مستوى أو درجة الأمن الصناعى ومعدل إصابة العمل لأمكن قياس العلاقــة بيــن المتغيرين بالإرتباط أو الإنحدار ولأمكن استخدام أسلوب تحليل الإنحدار لإختيار معنوية العلاقة بين درجة الأمن الصناعى وإصابة العمل. أما وأن القياســات ترتيبــية علــى هــذا النحو فلابد من وسيلة أخرى تلائم طبيعة القياسات كما أنه فى حالات أخرى فإنه بدافع السهولة وإمكانية التضحية ببعض الدقة فى سبيل تسهيل العمليات الحسابية وسرعة الوصول إلى قرار فإننا فى أى من الحالات الثلاث (عدم معرفة شكل التوزيع وطبيعة القياسات وأنها ترتيبية أو نوعية (أو وصفية) والسهولة والسرعة) فإننا نلجأ إلى أساليب إحصائية لا تعتمد علــى معــالم أو مؤشــرات أو بارامــترات تعرف باسم الأساليب الإحصائية اللامعلمية او اللابارامترية أو التى لا تعتمد على توزيع معين (Free or Non-parametric Distribution).

وهــذه الأســاليب وإن كانت مفيدة فى الحالات التى ذكرناها إلا أنه يعاب علــيها أنها تتجاهل قدراً من المعلومات التى قد تكون متاحة عن المتغير إذ أن اســتخدام الترتيــب قد يتغاضى عن الفرق فى القيمة بين مفردتين متتاليتين فى الترتيب مما يجعلها أقل حساسية للقيم القليلة المتطرفة ولكنها لهذا السبب عادة ما

تكـــون أقل كفاءة من الأساليب البارامترية. كما أنه فى الحالات التى يمكن فيها استخدام اختبار بارامترى (اختبارات مثلاً) وآخر لا بارامترى (اختبار الإشارة مـــثلاً) فإن الأخير عادة ما يكون أقل قوة من سابقه وهو ما سوف نشير إليه فى الفقـــرة التالـــية الخاصـــة بالقوة المكافئة كما سنشير إليه فى حينه عندما تتوافر شروط استخدام المدخلين فى الاختبار.

وهذه الأساليب قد تهدف إلى الوصف دون التحليل والإستنتاج ومنها الوسيط والمنوال ونصف المدى الربيعى ومعامل الإلتواء الربيعى ومعامل إرتباط الرتب ومعـــامل الـــتوافق وهو ما عرضنا له فى مرحلة سابقة. كما أنها قد تهدف إلى الإستنتاج الإحصائى وإتخاذ القرارات وهو ما يعنينا فى المرحلة الحالية. وسوف نعرض نماذج من الإختبارات الإحصائية اللابارامترية أو اللامعلمية وفى جميع الأحـــوال فإن خطوات الإختبار الإحصائى سوف تكون واحدة مهما اختلفت أداة الإختبار بارامترى أو لابارامترى.

القوة المكافئة للإختبار:

تعـــرف قوة الاختبار بأنها احتمال رفض الفرض العدمى وهو صحيح أى قـــدرة الإختبار على تجنب الخطأ من النوع الثانى أو حماية الباحث من الوقوع فى هذا الخطأ الذى يشار إليه بأنه خطأ من النوع الثانى.

وعـــادة مـــا يســـتخدم هـــذا المعـــيار للمقارنة بين الاختبارات الإحصائية اللابارامـــترية ونظائـــرها البارامترية. ولهذا الغرض سوف نعرف قاعدة القوة المكافئة للاختبار (Power-efficiency) على النحو التالى:

إذا كـــان حجـــم العينة اللازم لإجراء الأختبار أ هو $ن_أ$ وهو الحجم الذى يحقق للإختبار قوة معينة وكان يمكن الوصول إلى ذات الدرجة من القوة بإختبار آخر ب بعينة حجمها $ن_ب$ فإن:

القوة المكافئة للإختبار ب = $(ن_أ \div ن_ب) \times ١٠٠$

فمثلاً إذا كان لإجراء الأختبار أ بقوة معينة فإنه يلزم سحب عينة عشوائية حجمها ن$_{أ}$ = ٢٠ مفردة. وللوصول إلى ذات القوة باختبار آخر ب فإنه يلزم سحب عينة عشوائية حجمها ن$_{ب}$ = ٢٥ مفردة فإن:

القوة المكافئة للإختبار ب = $\frac{٢٠}{٢٥}$ × ١٠٠ = ٨٠٪

بمعـــنى أنـــنا نحـــتاج إلى ١٠٠ مفردة باستخدام الإختبار ب لتحقيق نفس الدرجة من القوة باستخدام ٨٠ وحدة فقط مع استخدام الإختبار أ.

وســـوف نـــتعرض فـــى بقية هذا الباب لأمثلة من الإختبارات الإحصائية اللابارامترية:

- ❀ باستخدام عينة واحدة.
- ❀ **باستخدام عينتين مستقلتين.**
- ❀ باستخدام عدة عينات مستقلة.

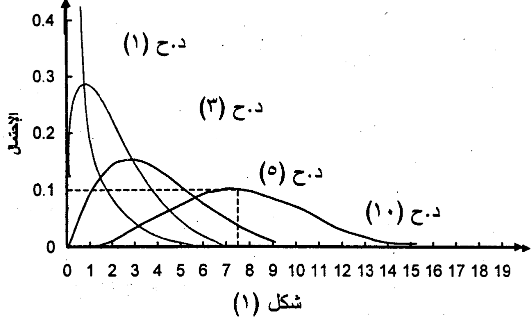
مـــع الإشـــارة إلـــى الإختـــبار البارامترى المناظر-إذا ما توافرت شروط استخدامه-وقو الإختبار اللابارامترى مقارناً بقوة الإختبار البارامترى المناظر.

وسوف نبدأ بتقديم توزيع احتمالى شائع الاستخدام فى الإختبارات اللامعلمية خاصة فى حالة البيانات التصنيفية وهو توزيع كا^{٢}.

(٣) توزيع كا٢ χ^2-Distribution:

توزيـــع كا^{٢} هو توزيع إحتمالى اكتشفه الإحصائيان المعروفان سير رونالد فيشر وكارل بيرسون R.A. Fisher & Karl Person فى أوائل القرن الحالى ولايـــتوقف شكل التوزيع على عدد مفردات الدراسة ولكن يتحدد شكل التوزيع على أساس عدد درجات الحرية وبالتالى فهو ليس توزيعاً وحيداً ولكنها مجموعة مـــن توزيعـــات تختلف باختلاف درجات الحرية. والشكل التالى يوضح صوراً مختلفة لهذا التوزيع لأعداد مختلفة من درجات الحرية ويتبين منه أن توزيع كا^{٢} يكون شديد الإلتواء إلى اليمين متى كان عدد درجات الحرية صغيراً ولكنه يأخذ

فى التماثل كلما زاد عدد درجات الحرية أنظر شكل (٥-١) التالى-ويعطى رقم () الإحـــتمالات أو المساحات تحت منحنى توزيع كا٢ Chi-Square بدرجات حرية مختلفة.



شكل (١)

توزيع كا٢ لدرجات حرية مختلفة

وأنـــه وإن كان هذا التوزيع قد عرف باستخدامه فى الإختبارات الإحصائية حيـــن تكون القراءات نوعية أو تصنيفية (Categorical) أو حين تتعدد النسب حيـــث يتعذر استخدام إختبارات النسبة التى تعرضنا لها فى فصل سابق إلا أنه يمكـــن استخدامه مع بقية أنواع القياسات الترتيبية أو بفترة أو نسبية متى أمكن رصد هذه المشاهدات عن المتغير مقابل فئات.

ويستخدم هذا التوزيع فى إجراء الإختبارات التالية:

١- اختـــبار جودة المطابقة أى مدى تشابه مشاهد لتوزيع نظرى وهو اختبار يعتمد على عينة واحدة.

٢- اختبار استقلال ظاهرتين وهو اختبار يعتمد أيضاً على عينة واحدة.

٣- اختـــبار تجانس توزيع عدة ظواهر فى مجتمع واحد وهو يعتمد على عدة عينات.

٤- الاختبارات الخاصة بالتباين σ^2 وتقديرها بفترة ثقة.

ولكننا سوف نقتصر فى المرحلة الحالية على اختبار جودة المطابقة واستقلال توزيع ظاهرتين فى مجتمع واحد واختبار تجانس توزيع ظاهرة ما فى عدة مجتمعات. مع إشارة عابرة إلى استخدام كا^2 فى الاختبارات المتعلقة بالتباين σ^2 وتقديرها بفترة ثقة.

أ: اختبار كا^2 لجودة المطابقة :

كثيراً ما يعنى الباحث بدراسة ما إذا كان التوزيع المشاهد لظاهرة ما فى أحد المجتمعات لا يختلف عن توزيع ظاهرى أو متوقع لهذه الظاهرة حيث يتحدد هذا التوزيع النظرى على أساس فروض معينة ، وإجراء اختبار كهذا فإنه يمكن مقارنة التكرارات المشاهدة (هـ$_ر$) للظاهرة موضوع الدراسة بالتكرارات النظرية أو المتوقعة (ت$_ر$) ، ر = ١ ، ٢ ، ...... ، ك تمثل الخلايا التى تتوزع داخلها التكرارات المشاهدة والنظرية وكلما اقترب التوزيع المشاهد من التوزيع النظرى كلما قلت قيمة الفروق فـ$_ر$ = هـ$_ر$ - ت$_ر$ وتنعدم الفروق تماماً عندما يتطابق التوزيعين المشاهد والنظرى وبالتالى فإنه يمكن أن تكون هذه الفروق أساساً لوسيلة أو أداة اختبار إحصائى ويمكن أن نثبت أن المتغير العشوائى:

$$\text{كا}^2 = \text{مجـ}_{ر} \frac{\text{ف}_{ر}^2}{\text{ت}_{ر}} = \text{مجـ}_{ر} \frac{(\text{هـ}_{ر} - \text{ت}_{ر})^2}{\text{ت}_{ر}} \qquad (1)$$

له توزيع كا^2 بدرجات حرية ك - ١

ويمكن استخدام قيم التوزيع الاحتمالى كا^2 بدرجات حرية ك-١ فى إجراء اختبار كهذا حيث ك : عدد الخلايا أو الفئات أو النواتج الممكنة للتجربة.

ويمكن تلخيص خطوات إجراء اختبار جودة المطابقة فى الأتى:

(١) نبدأ بتحديد الفرض المراد اختباره أى تحديد التوزيع النظرى للمتغير وتقدر القيم المتوقعة فى العينة العشوائية على أساس هذا الفرض النظرى ويرصد

مقابلهـــا القراءات المشاهدة ونرمز للتكرارات المشاهدة فى كل فئة أو خلية بالرمز $هـ_ر$ وللتكرارات المتوقعة $ت_ر$ أى

أنـــه سيكون لدينا ر من التكرارات المشاهدة هـــ يناظرها ر من التكرارات المتوقعة ت.

(٢) نحســب المقــدار: $كــا^٢ = مجـ_ر \frac{(هـ_ر - ت_ر)^٢}{ت_ر}$ والجمع بالنسبة لجميع الخلايا ك.

(٣) نحــدد مســتوى المعــنوية α منذ البداية (٥٪ أو ١٪) أو أى قيمة أخرى يختارها الباحث.

(٤) ونقــارن قيمة $كا^{٢*}$ (المحسوبة من البيانات) فى الخطــوة الســابقة بقيمــة $كــا^٢_{(ك-١، \alpha)}$ مــن الجــدول الخاص بتوزيع $كا^٢$ الموضح فى جدول (د) بــالملاحق. ونــرفض الفــرض الأصلى بعدم وجود اختلاف معنوى بين الــتوزيع المشاهد والتوزيع النظرى عند مستوى المعنوية α إذا كانت $كا^٢$ المحسوبة لقيمة $كا^٢$ من الجدول ويقبل الفرض العدمى فيما عدا ذلك. أى أن الاختبار فى هذه الحالة هو اختبار فى اتجاه واحد طرف أيمن.

وتوضح الأمثلة التالية طريقة إجراء هذا الاختبار.

مثال (١):

تعــد جــداول الأعــداد العشوائية بحيث تكون الأرقام من صفر إلى ٩ فى ترتيــب عشــوائــى فــى كل عدد ، وبحيث تكون فرص ظهور هذه الأرقام متساوية وتساوى ٠,١ بالنسبة لكل رقم.

وللــتحقق مــن توافــر هذه الخاصية فى تركيب جداول الأعداد العشوائية (جــدول رقــم ٧) بــالملاحق ، فقد أخذت مفردات العمودين الأول والثانى من الصــفحة الثانية من الجدول وكانت نتيجة حصر التكرارات المشاهدة لكل رقم وكذلك التكرارات المقدمة لكل كما يوضحها الجدول التالى جدول (٥-٢). اختبر

فرض جودة المطابقة – أى مطابقة تركيب الأعداد العشوائية لخاصية إعداد تلك الجداول عند مستوى المعنوية ٥٪.

الحل :

جدول (٢)

التوزيع المشاهد والنظرى (المتوقع)

للأرقام بالأعداد العشوائية

الرقم	التكرار المشاهد (هـ ر)		التكرار المتوقع (ت ر)	كا٢
صفر	٢٦	٠,١٠	٢٥	٠,٠٤
١	٢١	٠,١٠	٢٥	٠,٠٤
٢	٢٥	٠,١٠	٢٥	صفر
٣	٢٥	٠,١٠	٢٥	صفر
٤	٢٨	٠,١٠	٢٥	٠,٣٦
٥	٢٧	٠,١٠	٢٥	٠,١٦
٦	٢١	٠,١٠	٢٥	٠,٦٤
٧	٢٩	٠,١٠	٢٥	٠,٦٤
٨	٢٣	٠,١٠	٢٥	٠,١٦
٩	٢٥	٠,١٠	٢٥	صفر
مجموع	٢٥٠		٢٥	٢,٦٤

ويتم اختبار فرض جودة المطابقة على النحو التالى:

(١) الفرض : أن جـداول الأعـداد العشـوائية – الجزء الذى تم اسـتخدامه فى الاختبار – تتوافر له خاصية الجداول العشوائية.

(٢) مستوى المعنوية : α = ٥٪

(٣) أداة الاختبار : $كا^{٢} = \frac{(هـ_{ر} - ت_{ر})^{٢}}{ت_{ر}}$

يــرفض الفــرض العدمــى (وهو الفرض بعدم اختلاف التوزيع المشاهد عــن الــتوزيع الــنظــرى أو المتوقــع) إذا كانت كا^{٢*} ≤ كا^{٢} (١٠-١ ، ٥٪) من الجدول = ١٦,٩٠٩ .

(٤) ومن البيانات : كا^{٢*} = ٢,٦٤ < ١٦,٩١٩

∴ يقــبل فرض العدم أى أن جداول الأعداد العشوائية يتفق فى تركيبه مع خاصية الجداول العشوائية.

مثال (٣):

الجــدول التالى يوضح التوزيع التكرارى لنتائج اختبار ما طبق على عينة عشــوائية مكونـــة من ٨٦ عاملاً فى نهاية برنامج تدريبى. اختبر الفرض بأن توزيــع درجــات الاختبار تتبع توزيعاً معتاداً متوسطه وتباينه هو نفس متوسط وتباين توزيع العينة وذلك عند مستوى المعنوية ٥٪ .

ملحوظة : لــم تحدد قيم المؤشرات التى تعين التوزيع المعتاد النظرى فى هذا المــثال ولذلك تعين علينا تقديرها من بينات العينة وهذه المؤشرات هى $\mu = \overline{س}$ = ٢٦,١ درجة ، σ = ع = ٦,٤٥ درجة.

الحل :

جدول (٣)

التوزيع النظرى والمشاهد للدرجات فى برنامج تدريبى لعدد ٨٦ عاملاً

الفئات	التكرارت المشاهدة هـ ر		الحدود العليا س	$\frac{س - \overline{س}}{ع}$ = ى	الاحتمالات التجميعية	التكرارات ت ر		كا٢ ر
٥-٧	صفر		٧,٥	-٢,٨٨	٠,٠٠٢	٠,٢		
٨-١٠	صفر		١٠,٥	-٢,٤٢	٠,٠٠٨	٠,٢		
١١-١٣	٤	٧	١٣,٥	-١,٩٥	٠,٠١٦	١,٥	٥,٨	٠,٢٤٨
١٤-١٦	٣		١٦,٥	-١,٤٩	٠,٠٦٨	٣,٦		
١٧-١٩	٨		١٩,٥	-١,٠٢	٠,١٥٤	٧,٤		٠,٤٩
٢٠-٢٢	١٣		٢٢,٥	-٠,٥٦	٠,٢٨٨	١١,٦		٠,١٦٩
٢٣-٢٥	٩		٢٥,٥	-٠,٠٩	٠,٤٦٤	١٥,١		٢,٤٦٤
٢٦-٢٨	١٧		٢٨,٥	٠,٣٧	٠,٨٠٠	١٥,٥		٠,١٤٥
٢٩-٣١	١		٣١,٥	٠,٨٤	٠,٩٠٣	١٣,٤		٠,٠٢٧
٣٢-٣٤	٨		٣٤,٥	١,٣٠	٠,٩٦٢	٨,٩		٠,٠٩١
٣٥-٣٧	٥		٣٧,٥	١,٧٧	٠,٩٨٧	٥,٠		
٣٨-٤٠	٤		٤٠,٥	٢,٢٣	٠,٩٩٧	٣,٢		
٤١-٤٣	١	١٠	٤٣,٥	٢,٧٠	٠,٩٩٩	٠,٨	٨,٣	٠,٣٤٨
٤٤-٤٦	صفر		٤٦,٥	٢,١٦	١,٠٠٠	٠,٢		
٤٧ <	صفر							
	٨٦							٣,٥٤١

وبمقارنـــة كــا٢* = ٣,٥٤١ بقيمة كا٢ = ٨ ـ ١ ـ ٢ = ٥ ، ٠,٠٥ من الجــدول = ١١,٠٧٠ نجــد أن كــا٢* > كا٢ من الجدول وبالتالى يقبل الفرض الأصــلى بــأن الــتوزيع التكرارى المشاهــد يتبع توزيعــاً معتــاداً متوسطه μ = ٢٦,١ وتباينه = $(٦,٤٥)^{٢}$ ويلاحظ فى هذا المثال أن:

(١) التكرارات المتوقعة $ت_ر$ فى الفئات الأربعة الأولى وفى كل من الخمسة الأخيرة تقل عن ٥ تكرارات فى كل فئة لذلك فقد أدمجت الفئات الأربع الأولى وكذلك الخمس الأخيرة وجمعت تكراراتها المتوقعة والمشاهدة لتحقيق التقارب لتوزيع كا٢ وهذه قاعدة عملية شائعة الستخدام ويتعين استخدامها حينما تقل التكرارات المتوقعة فى خلية أو أكثر عن ٥ وحدات ثم تحسب قيمة كا٢ بعد إجراء هذا الإدماج أى أن كا٢ فى الفئات الأربع الأولى (٥-١٦)

$$= \frac{(٧ - ٥,٨)^٢}{٥,٨} = ٠,٢٤٨$$

وبالمثل فإن كا٢ فى الخمس الأخيرة ٣٥ فأكثر $= \frac{(١٠ - ٨,٣)^٢}{٨,٣} = ٠,٣٤٨$

(٢) إن المتغير العشوائى $\frac{س - \bar{س}}{ع} = \frac{س - \mu}{\sigma}$ له توزيع معتاد معيارى (ى) متوسطه الصفر وتباينه الوحدة وبالتالى فإنه يمكن إيجاد المساحة المحصورة بين $-\infty$ وأى قيمة (ى) على المحور الأفقى باستخدام جداول التوزيع المعتاد المعيارى.

فمثلاً :

ح ($-\infty$ إلى $-٢,٨٢$) = ٠,٠٠٢

ح ($-\infty$ إلى $-٢,٤٢$) = ٠,٠٠٨

ح ($-\infty$ إلى $-١,٩٥$) = ٠,٠١٦

وتكون المساحة المناظرة للفئة ٥ – ٧ = احتمال وقوع مشاهدة فى هذه الفترة = ٠,٠٠٢

$\therefore$ التكرار المتوقع $= \frac{٨٦ \times ٢}{١٠٠٠} = ٠,٢$

والمساحة المناظرة للفئة ٨ – ١٠ = احتمال وقوع مشاهدة فى هذه الفترة ٠,٠٠٨ – ٠,٠٠٢ = ٠,٠٠٦

ويكون التكرار المتوقع = $\frac{٨٦ \times ٢}{١٠٠٠}$ = ٠,٥ .... وهكذا.

(٣) عــدد درجــات الحــرية لقيمة كا^{٢} = عدد الفئات (أو الخلايا) – ١ – عدد المؤشرات أو المعالم التى حسبت من العينة. وفى مثالنا هذا درجات الحرية = عدد الخلايا بعد إدماج الأربعة الأولى والخمس الأخيرة – ١ – درجتين مقابل تقدير المتوسط الحقيقى والتباين الحقيقى = ٨ – ١ – ٢ = ٥ .

ملحوظة:

(١) إذا أجــرى اختــبار يعتمد على توزيع كا^{٢} بدرجة حرية واحدة أى كانت القراءات المشاهدة والمتوقعة تقع فى خليتين اثنين فقط فإنه يتعين أن تكون القــراءات المــتوقعة فــى كل خليــة لا تقل عن خمسة. وإذا كان عدد الخلايـــــا ك>٢ فإنه يتعين ألا يزيد عدد الخلايا التى تقل فيها القراءات المتوقعة عن٥ عن٢٠٪ من مجموع عدد الخلايا أو كان التكرار المتوقع فى إحداها يقل عن١فإنه يمكن إدماج الخلايا المتعاقبة كمافى المثال(٥-٢)

(٢) وغــنى عــن الذكر أن كا^{٢}* دائماً موجبة القيمة أما إذا كانت قيمة كا^{٢}*= صــفر فــإن ذلك يستوجب إعادة فحص البيانات وكيف جمعت فقد تكون الطــريقة التى صنفت البيانات على أساسها قد أدت إلى عدم ظهور فروقاً حقيقية بين القراءات المشاهدة والمتوقعة.

استخدام توزيع كا^{٢} لاختبار جودة المطابقة لتوزيع غير مستمر متعدد النواتج:

كما يمكن استخدام توزيع كا^{٢} فى اختبار التجارب العشوائية متعددة النواتج أى ك ، ك > ٢ (multinomial). فــتجربة إلقــاء زهرة النرد ينتج عنها ٦ نواتج متماثلة مستقلة وثابتة عند تكرار التجربة (النواتج هى٣،٢،١ ،٤ ،٥ ،٦). وبالمثل فإن استطلاع الرأى فى موضوع ما يختلف من الموافقة إلى الرفض إلى الإمتــناع عــن إبداء الرأى وهى أيضاً نواتج متماثلة مستقلة وثابتة وفى جميع

الأحـــوال فـــإن $ل_١ + ل_٢ + ل_٣ + ... + ل_ك = ١$. وفـــى حـــالات كهذه فإن الإختـــبارات التى تعتمد على توزيع ذو الحدين أو الطبيعى المعيارى لا تصلح للاستخدام والبديل هنا هو توزيع $كا^٢$ وهو ما يوضحه المثال التالى:

مثال (٣):

أجريـــت دراســـة تسويقية للتعرف على مدى تفضيل المستهلكين لسلعة ما (المـــنظفات الصناعية مثلاً) من أنواع مختلفة عددها ٥ أنواع فقد سحبت عينة عشوائية مكونة من ١٠٠٠ من المشترين لهذه السلعة وكان عدد من فضلوا نوعاً على الأنواع الأخرى كالآتى:

النوع	أ	ب	جـ	د	هـ	المجموع
عدد الذين فضلوا هذا النوع على غيره	٢١٤	٢٣١	١٨٢	١٥٤	٢١٩	١٠٠٠

عـــند مســـتوى المعنوية ١٪ اختبر ما إذا كانت نسبة من فضلوا نوعاً على الآخر لا تختلف من نوع لآخر.

الحل :

(١) الفرض العدمى هو أن $ل_١ = ل_٢ = ل_٣ = ل_٤ = ل_٥$

$$حيث\ ل_ر = \frac{ن_ر}{ن} = ٠,٢٠$$

الفرض البديل هو : أن احتمال التفضيل لنوع على الآخر $\ne$ ٠,٢٠ لنوعين من الأنواع الخمسة على الأقل.

(٢) أداة الاختبار هى:

$$كا^{٢*} = مجـ \frac{(هـ - ت)^٢}{ت}$$ لها توزيع :

$كا^٢$ بدرجات حرية (٥ – ١ = ٤) = ١٣,٢٧٧

ويرفض الفرض العدمى إذا كانت $كا^{٢*}$ > ١٣,٢٧٧

(٣) ومن البيانات :

$$\text{كا}^{٢*} = \frac{(٢١٤ - ٢٠٠)^{٢}}{٢٠٠} + \frac{(٢٣١ - ٢٠٠)^{٢}}{٢٠٠} + \frac{(١٨٢ - ٢٠٠)^{٢}}{٢٠٠}$$

$$+ \frac{(١٥٤ - ٢٠٠)^{٢}}{٢٠٠} + \frac{(٢١٩ - ٢٠٠)^{٢}}{٢٠٠}$$

$$= ٠,٩٨٠ + ٤,٨٠٥ + ١,٦٢٠ + ١٨,٥٨٠ + ١,٨٠٥$$

$$= ١٩,٧٩٠$$

(٤) القرار : $\text{كا}^{٢*} = ١٩,٧٩٠ > ١٣,٢٧٧$

∴ يرفض الفرض العدمى ويقبل الفرض البديل.

وهذا يلاحظ أن الجزء الأكبر من اختلاف كا²* عن كا² يرجع إلى اختلاف معنوى فى نسبة الذين فضلوا النوع ب ، د عن بقية الأنواع.

وهذه النسبة = $\frac{(٤,٨٠٥ - ١٠,٥٨٠)}{٢٠٠} \times ١٠٠ = ٧٧,٧٤\%$

ويفيد مثل هذا التحليل فى اتخاذ قرار تسويقى مناسب بالنسبة لمبيعات هذه الأنواع الخمسة.

ب : اختبار كا² لإستقلال توزيع ظاهرتين أو متغيرين فى مجتمع واحد:

يهـــدف هذا الاختبار إلى اختبار استقلال توزيع متغيرين فى مجتمع واحد ، كالمستوى التعليمى لرب الأسرة ودخله السنوى أو حجم الأسرة أو عدد الأطفال والمستوى التعليمى للأب أو للأم.

وقد تكون قياسات أحد المتغيرين أو كلاهما وصفية.

وتخـــتار عينة عشوائية واحدة حجمها ن ثم تصنف مفردات العينة بحسب فئات كل من المتغيرين فى جداول توافق مكون من ل صف ، م عمود.

<u>خطوات الاختبار:</u>

(١) يجـــرى اختبار توزيع المتغيرين بتقدير التكرارات المتوقعة فى كل خلية من خلايا جدول التوافق (Contingency table) وعددها ل × م خلية.

وإذا رمزنا إلى التكرارات المشاهدة فى كل خلية بالرمز $هـ_{ور}$ وللتكرارات المتوقعة بالرمز $ت_{ور}$ حيث ر عدد الصفوف ، و عدد الأعمدة ، (ر = ١ ، ٢ ، ... ، ل ، و = ١ ، ٢ ، ... ، م) وحيث تتحدد $ت_{ور}$ فى حالة صحة فرض استقلال المتغيرين كالآتى:

$$ت_{ور} = ن (\frac{ك_{ر.}}{ن} \frac{ك_{.و}}{ن}) = (\frac{ك_{ر.} \times ك_{.و}}{ن}) \qquad (٢)$$

= $\frac{\text{مجموع الصف الذى تقع القراءات المشاهدة} \times \text{مجموع العمود الذى تقع فيه هذه القراءة}}{\text{المجموع الكلى}}$

وذلك تطبيقاً لقاعدة ضرب الاحتمالات فى حالة الاستقلال.

(٢) وإذا صح فرض الاستقلال فإن المتغير العشوائى:

$$كا^{٢*} = مجـ_{ور} \frac{(هـ_{ور} - ت_{ور})^{٢}}{ت_{ور}}$$

له توزيع $كا^{٢}$ بدرجات حرية (ل-١) (م-١) (٣)

(٣) تحسب $كا^{٢*}$ من بيانات العينة وتقارن بقيمة $كا^{٢}_{(ل-١م-١،\alpha)}$ ويقبل أو يرفض فرض الاستقلال حسب الأحوال.

والمثال التالى يوضح هذا الأسلوب:

مثال (٤):

سحبت عينة عشوائية مكونة من ٤٠٠ شخصاً ثم صنفت إجاباتهم عن المستوى التعليمى لكل منهم وجملة دخله السنوى فكانت النتائج كما يوضحها الجدول التالى :

اختبر الفرض باستقلال الدخل السنوى عن المستوى التعليمى عند مستوى المعنوية ٥٪ .

جدول (٤)

توزيع عينة من ٤٠٠ فرداً بحسب الدخل السنوى والمستوى التعليمى

الدخل السنوى أ	المستوى التعليمى ب						المجموع
	المرحلة الثانوية		المرحلة الجامعية		الدراسات العليا		
	$ر_{هـ١}$	$ر_{ت١}$	$ر_{هـ٢}$	$ر_{ت٢}$	$ر_{هـ٣}$	$ر_{ت٣}$	
منخفض	٦٣	٥١,٠٠	٧٨	٧٦,٥٠	٢٩	٤٢,٥٠	١٧٠
متوسط	٤٨	٥٢,٥٠	٨٢	٧٨,٧٥	٤٥	٤٣,٧٥	١٧٥
مرتفع	٩	١٦,٥	٢٠	٢٤,٧٥	٢٦	١٣,٧٥	٥٥
المجموع	١٢٠	١٢٠	١٨٠	١٨٠	١٠٠	١٠٠	٤٠٠

الملاحــظ فــى هذا المثال أن القياسات تعتمد على العد وليس على القياس الكمى وعليه فلم يكن يصلح لاختبار فرض كهذا أى من الاختبارات البارامترية (التى تفترض توزيعاً معيناً له معالم أو بارامترات أو مؤشرات ذا تقيم معلومة).

ولإجــراء اختبار استقلال المستوى التعليمى عن الدخل سوف نستخدم أداة الاختبار:

$كا^٢ = \frac{مجـ(هـ - ت)^٢}{ت}$ حيــث يــتم حساب التكرارات لجميع الخلايا وعددها ل × م ويتم الجمع لجميع الخلايا. ولحساب التكرار المتوقع فى كل خلية فإننا نلاحظ أنه إذا صح الفرض باستقلال توزيع الظاهرتين فإن:

ح (إن يكون الشخص منخفض الدخل وقد أتم المرحلة الثانوية فى التعليم)

= $ح_١ × ح_٢$ = ح ($أ_{١١}$: منخفض الدخل) × ح ($ب_{١١}$: أتم المرحلة الثانوية).

$$= \frac{١٧٠}{٤٠٠} × \frac{١٢٠}{٤٠٠}$$

∴ $ت_{١١}$: العدد المتوقع للأشخاص الذين هم من ذوى الدخل المنخفض من الذين أتموا التعليم الثانوى =

$$ن \times ح(أ_{١١})\ ح(ب_{١١}) = ن \times \frac{\text{مجموع الصف الأول} \times \text{مجموع العمود الأول}}{\text{المجموع الكلى}}$$

$$= ٤٠٠ \quad \frac{١٧٠}{٤٠٠} \times \frac{١٢٠}{٤٠٠} = \frac{١٧٠ \times ١٢٠}{٤٠٠} = ٥١ \text{ شخصاً}$$

وبالمثل بالنسبة لباقى الخلايا وتكون التكرارات المتوقعة لجميع خلايا الجدول كما هو الحال بالجدول (٤) السابق.

ويستكمل الاختبار كالآتى:

(١) الفرض الأصلى : إن الدخل السنوى مستقل عن المستوى التعليمى والفرض البديل إنهما غير مستقلين.

(٢) مستوى المعنوية α = ٥٪

(٣) وسيلة الاختبار $كا^{٢} = \frac{مجـ(هـ - ت)^{٢}}{ت}$

والجمع بالنسبة لجميع الخلايا فى جميع الصفوف وجميع الأعمدة. ويرفض الفرض الأصلى إذا كانت أى المحسوبة $\geq كا^{٢}$ (ل-١) (م-١) ، α).

(٤) ومن البيانات فإن :

$$كا^{٢*} = \frac{(٦٣-٥١)^{٢}}{٥١} + \frac{(٤٨-٥٢,٥)^{٢}}{٥٢,٥} + \frac{(٩ - ١٦,٥)^{٢}}{١٦,٥} +$$

$$\frac{(٧٨-٧٦,٥)^{٢}}{٤٢,٥} + \frac{(٨٢-٧٨,٧٥)^{٢}}{٧٨,٧٥} + \frac{(٢٠ - ٢٤,٧٥)^{٢}}{٢٤,٧٥} +$$

$$\frac{(٢٩-٤٢,٥)^{٢}}{٤٢,٥} + \frac{(٤٥-٤٣,٧٥)^{٢}}{٤٣,٧٥} + \frac{(٢٦ - ١٣,٧٥)^{٢}}{١٣,٧٥} +$$

= ٢,٢٨ + ٠,٣٨ + ٣,٤١ + ٠,٠٣ + ٠,٩٢

+ ٤,٢٩ + ٠,٢٤ + ١٠,٩١ = ٢٢,٣٩

ولما كانت $كا^{٢}_{٤\,،\,٥٪}$ = ٩,٤٨٨ < $كا^{٢*}$ = ٢٢,٣٩ لذلك يرفض الفرض باستقلال توزيع الدخل السنوى عن المستوى التعليمى لصاحب هذا الدخل عند مستوى المعنوية ٥٪ .

ملحوظة:

يمكن من ملاحظة القيم التى تكون فى مجموعها كا² المحسوبة ومعرفة أيهما أكبر مساهمة فى تكوين هذا المجموع وبالتالى يمكن معرفة أى فئات الدخل أكثر من غيرها إعتماداً على المستوى التعليمى وتطبيقاً لذلك فى المثال السابق يتضح أن الدخل المرتفع قد أسهم فى قيمة كا² بـــ ٢,٢٨+٠,٩٢+١٠,٩١ = ١٤,١١ أى حوالى ٦٣٪ مما يعنى أن ارتفاع الدخل له أثر معنوى فى تحقيق مستوى تعليمى مرتفع.

حالة خاصة : اختبار كا² : جدول ٢ × ٢ :

إذا كانت البيانات مبوبة فى جدول مزدوج مكون من صفين وعمودين فإنه يمكن الوصول إلى قيمة كا²* بالطريقة التالية بدلاً من الطريقة العامة التى شرحناها فى الفقرات السابقة.

الخاصية (٢)	الخاصية (١)		المجموع
	الفئة (١/١)	الفئة (٢/١)	
الفئة (١/٢)	أ	ب	أ + ب
الفئة (٢/٢)	جـ	د	جـ + د
المجموع	أ + جـ	ب + د	ن

$$\text{كا}^2 \text{ المحسوبة} = \frac{(\text{أد} - \text{ب جـ})^2 \text{ن}}{(\text{أ} + \text{جـ})(\text{ب} + \text{د})(\text{أ} + \text{ب})(\text{جـ} + \text{د})} \qquad (٤)$$

مثال (٥):

الجدول التالى رقم (٥) يوضح توزيع مفردات عينة عشوائية من ٤١٢ فرداً من الريف والحضر بحسب موافقتهم أو رفضهم لمشروع التأمينات الإجتماعية. اختبر عند مستوى المعنوية ٥٪ استقلال الرأى عن محل الإقامة.

الحل :

جدول (٥)

توزيع عينة من ٤١٢ فرداً حسب محل الإقامة الرأى فى التأمينات الإجتماعية

الإقامة	الرأى		المجموع
	موافق	معارض	
ريف	٨٤	١١١	١٩٥
حضر	١٢٢	٩٥	٢١٧
المجموع	٢٠٦	٢٠٦	٤١٣

$$كا^{٢*} = \frac{(٨٤×٩٥ - ١٢٢ × ١١١)^٢}{٢٠٦ × ٢٠٦× ٢١٧×١٩٥} = ٧,١٠$$

وعند مستوى المعنوية ٥٪ فإن كا^٢ (٢-١) (٢-١) من الجدول = ٣,٨٤

أى أن كا^{٢*} = ٧,١٠ > كا^٢ من الجدول = ٣,٨٤

لذلك يرفض الفرض بأنه لا فرق بين تقبل أفراد الريف والحضر للمشروع عند مستوى المعنوية ٥٪ .

وكـــان يمكن حل نفس المثال باستخدام الطريقة العامة أى حساب $ت_{ور}$ ثم اســتخدام وسيلة الاختبار كا^٢ = مجـ $\frac{(هـ - ت)^٢}{ت}$ مع الجمع بالنسبة لجميع الخلايا وكنا سنحصل على نفس القيمة وبالتالى نصل إلى نفس القرار.

تصحيح ييتس :

يســتخدم تصحيح ييتس لتصحيح قيمة كا^٢ المحسوبة لمقارنتها بـــ كا^٢ من الجـــدول وتتضـــح أهمية هذا التصحيح إذا ما لاحظنا أن توزيع كا^٢ هو توزيع لمتغير مستمر بينما المتغير الذى تحسب من بياناته كا^٢ متغير متقطع. ويستخدم هـــذا التصحيح عندما تكون درجات الحرية لـــ كا^٢ درجة واحدة وذلك إذا كان

التوزيع مكونة من خليتين فقط تقع فيها التكرارات المشاهدة والمتوقعة وذلك فى حالـــة إختبار جودة المطابقة ، وهى حالة مطابقة التوزيع المشاهد لتوزيع ذى حدين. كما يكون اختبار $\text{كا}^{٢}$ بدرجة حرية واحدة إذا كنا بصدد جدول ٢ × ٢ فى اختبارات الاستقلال والتجانس.

وفى الحالة الأولى يتم التصحيح بطرح ٠,٥ من الفرق المطلق بين التكرار المشـــاهد والمـــتوقع فى كل من الخليتين. أى أن $\text{كا}^{٢}$ فى هذه الحالة تكون على الصورة.

$$\text{كا}^{٢}\ (\text{المحسوبة}) = \frac{(\text{أهـ}_{١}\text{−ت}_{١}\text{أ} - ٠,٥)^{٢}}{\text{ت}_{١}} + \frac{(\text{أهـ}_{٢}\text{−ت}_{٢}\text{أ} - ٠,٥)^{٢}}{\text{ت}_{٢}} \qquad (٥)$$

أما فى حالة جدول ٢ × ٢ فإننا نحصل على الصيغة المصححة كالآتى: ٠,٥

$$\text{كا}^{٢}\ (\text{المصححة}) = \frac{(\text{أد − ب جـ أ} - \frac{\text{ن}}{٢})\text{ن}}{(\text{أ + جـ})(\text{ب + د})(\text{أ + ب})(\text{جـ + د})} \qquad (٦)$$

وقـــد يؤثـــر هذا التصحيح على القرار أو رفض الفرض الأصلى إذا كانت قـــيمة كـــا٢ المحســـوبة بدون تصحيح قريبة من القيمة المعنوية التى تؤدى إلى رفض الفرض الأصلى. إذ أن التصحيح قد يؤدى إلى خفض قيمة كا٢ المحسوبة وبالـــتالى إلـــى تغيـــير القـــرار من الرفض إلى القبول أما إذا كانت قيمة كا٢ المحســـوبة لا تؤدى إلى رفض الفرض الأصلى أو أن قيمتها تختلف كثيراً عن قيمة كا٢ من الجدول فإن هذا التصحيح يفقد قيمته العملية.

مثال (٦):

يوضـــح الجدول التالى رقم (٦) نتائج دراسة أجريت على عينة مكونة من ٤٥ فـــرداً وتفضيلهم الأنباء المذاعة من الإذاعة على المنشورة بالصحف اختبر عند مستوى ٥٪ استقلال الرأى عن نوع الفرد.

جدول (٦)

توزيع أفراد عينة مكونة من ٤٥ فرداً بحسب النوع والرأى فى موضوع ما

الإجابة	الـــــــــــــــرأى		المجموع
	ذكور	إناث	
نعم	٢١	١٠	٣١
لا	٤	١٠	١٤
المجموع	٢٥	٢٠	٤٥

الحل :

من البيانات الموضحة بالجدول السابق يتبين أن:

$$\text{كا}^{٢}\ (\text{المحسوبة}) = \frac{(٢١\times١٠ - ٤\times١٠)^{٢}\ ٤٥}{٢٥\times٢٠\times١٤\times٣١} = ٥,٩٩$$

$$\text{كا}^{٢}\ (\text{المصححة}) = \frac{(|٢١\times١٠ - ٤\times١٠| - \frac{٤٥}{٢})^{٢}\ ٤٥}{٢٥\times٢٠\times١٤\times٣١} = ٤,٥١$$

وقيمة كا^{٢} المحسوبة تتجاوز قيمة كا^{٢} من الجدول بدرجة حريـــة واحدة عند α = ٥٪ وتســاوى ٣,٨٤ إذا حســبت كا^{٢} بدون تصحيح كما أن لتصحيح لا يؤدى إلى تغيير فى القرار ولم يكن من المحتمل أصلاً أن تقل القيمة المحسوبة بعد التصحيح عن ٣,٨٤ لذلك كان يمكن أن يصرف النظر عن التصحيح.

جـ: اختبار كا^{٢} لتجانس توزيع ظاهرة ما فى عدة مجتمعات :

لا تخـــتلف طـــريقة إجــراء اختــبار كهذا عن الإختبار الخاص باستقلال ظاهرتيــن باســتخدام توزيع كا^{٢} ولكنهما يختلفان من حيث طريقة بناء النموذج وتصـــميم الـــتجربة وبالـــتالى الفـــرض موضوع الاختبار. ففى حالة استقلال ظاهرتيــن فــإن حجــم العينة مثبت أو بمعنى آخر فإننا نختار عينة واحدة من

المجتمع ويتم تصنيف مفرداتها بحسب متغيرين أو صفتين وعليه فإن المجاميع الهامشية للصفوف والأعمدة غير مثبتة أما في حالة اختبار التجانس (Homogeneity) أي عدم اختلاف توزيع المتغير في عدة مجتمعات فيتم سحب عدة عينات من كل مجتمع عينة واحدة ذات حجم محدد ويتم رصد التوزيع المشاهد للمتغير في كل عينة وبالتالي فإن المجاميع الهامشية مثبتة في إتجاه واحد هو مجموع كل عينة.

خطوات الاختبار :

(١) الفرض العدمي: هو أنه لا يوجد اختلاف في توزيع المتغير في المجتمعات المختلفة أي أن التوزيع متجانس في المجتمعات المختلفة أما الفرض البديل: فهو أن التوزيع يختلف من مجتمع لآخر.

(٢) يحدد مستوى المعنوية α .

$$كا^{٢} = مجـ_{ور} \frac{(هـ - ت)^{٢}}{ت}$$

حيث هـ تمثل القراءات المشاهدة أي توزيع كل عينة على فئات المتغير ، ت تمثل التكرارات المتوقعة لكل تكرار مشاهد وتحسب كالمعتاد.

$$ت_{ور} = \frac{مجموع\ الصف\ (ر) \times مجموع\ العمود\ (و)}{ن}$$

لجميع ر = ١ ، ٢ ، ... ، ل ، و = ١ ، ٢ ، ... ، م ، ن : الحجم الكلي للعينات.

(٤) وإذا صح الفرض العدمي فإن المتغير العشوائي:

$$كا^{٢*} = مجـ_{ور} \frac{(هـ - ت)^{٢}}{ت}$$

له توزيع $كا^{٢}$ بدرجات حرية (ل-١) (م-١)

ويقبل أو يرفض الفرض العدمي بمقارنة $كا^{٢*}$ بـ $كا^{٢}$ (ل-١) (م-١).

مثال (٧):

يبيــن الجــدول (٧) توزيع من اجتازوا الاختبار فى نهاية برنامج تدريبى موحــد طبق فى ثلاثة أقسام مختلفة من إحدى المؤسسات. اختبر الفرض القائل بأن قدرات المتدربين متقاربة فى الأقسام الثلاث عند مستوى المعنوية ٥٪ .

جدول (٧)

توزيع التقديرات فى برنامج تدريبى فى عينات من ثلاث أقسام مختلفة

القسم	نتيجــــة الاختبـــــار				المجموع
	نجـــــاح		فشـــــل		
	ر هـ ١	ر ت ١	ر هـ ٢	ر ت ٢	
الأول	٥٠	٤٦,٧٥	٥	٨,٢٥	٥٥
الثانى	٤٧	٥١,٥٨	١٤	٩,١٥	٦١
الثالث	٥٦	٥٤,٦٧	٨	٩,٦	٦٤
المجموع	١٥٣	١٥٣	٢٧	٢٧	١٨٠

وتحسب التكرارات المتوقعة كالآتى:

$$\text{ت}_{١١} = \frac{١٥٣ \times ٥٥}{١٨٠} = ٤٦,٧٥$$

$$\text{ت}_{١٢} = \frac{١٥٣ \times ٦١}{١٨٠} = ٥١,٥٨$$

$$\text{ت}_{١٣} = \frac{١٥٣ \times ٦٤}{١٨٠} = ٥٤,٦٧$$

$$\text{ت}_{٢١} = \frac{٢٧ \times ٥٥}{١٨٠} = ٥٥ - ٤٦,٧٥ = ٨,٢٥$$

$$\text{ت}_{٢٢} = \frac{٢٧ \times ٦١}{١٨٠} = ٦١ - ٥١,٥٨ = ٩,١٥$$

$$\text{ت}_{٢٣} = \frac{٢٧ \times ٦٤}{١٨٠} = ٦٤ - ٥٤,٤ = ٩,٦٠$$

ويتم استكمال خطوات الاختبار كالآتى:

(١) الفرض الأصلى: أنــه فى ضوء نتيجة الاختبار فإن قدرات المتدربين لا تختلف معنوياً أى أن قدراتهم متجانسة فى الأقسام الثلاثة.

(٢) مستوى المعنوية ٥٪ .

(٣) وسيلة الاختبار كا^{٢} المحسوبة = $\text{مجـ}_{\text{ور}} \frac{(\text{هـ} - \text{ت})^{٢}}{\text{ت}}$

ويرفض الفرض الأصلى إذا كانت كا^{٢} المحسوبة ≤ كا^{٢} من الجدول عند درجات الحرية (ل-١) (م-١) ومستوى المعنوية ٥٪ .

(٤) ومن البيانات يتضح أن:

$$\text{كا}^{٢*} = \frac{(\text{٥٠}-\text{٤٦,٧٥})^{٢}}{\text{٤٦,٧٥}} + \frac{(\text{٤٧}-\text{٥١,٨٥})^{٢}}{\text{٥١,٨٥}} + \frac{(\text{٥٦}-\text{٥٤,٦٧})^{٢}}{\text{٥٤,٦٧}}$$

$$+ \frac{(\text{٥}-\text{٨,٢٥})^{٢}}{\text{٨,٢٥}} + \frac{(\text{١٤}-\text{٩,١٥})^{٢}}{\text{٩,١٥}} + \frac{(\text{٨}-\text{٩,٦})^{٢}}{\text{٩,٦}}$$

وهى أصغر من ٥,٩٩ = كا$^{٢}_{(\text{٢ ، ٠,٠٥})}$.

لذلــك يقــبل الفرض العدمى بعدم وجود اختلاف معنوى فى توزيع قدرات المتدربين فى الأقسام الثلاث عن مستوى المعنوية ٥٪ .

ومــن البديهى أننا وقد قبلنا الفرض العدمى فلا محل لتحليل مكونات كا^{٢*} لــتحديد أى الأقسام أسهمت أكثر من غيرها فى تكوين المجموع لــ كا^{٢*} . ولو أنــنا رفضنا الفرض بتجانس توزيع الدرجات فى المجتمعات الثلاث لكان علينا محاولــة تفســير درجــة إختلاف المجتمعات الثلاث بالرجوع إلى قيمة الحدود المختلفة التى تكون فى مجموعها القيمة الكلية لــ كا^{٢*} .

<u>ملاحظات ختامية بشأن استخدام كا^{٢} فى الاختبارات اللامعلمية:</u>

(١) إن اســتخدام كــا^{٢} فى الاختبارات الإحصائية الخاصة بجودة المطابقة أو اســتقلال توزيــع ظاهرتيــن لا يفترض شكلاً معيناً لتوزيع الظاهرة أو الظاهرتين فى المجتمع الأصــلى ولذلك يطلق عليه اختبار التوزيع الحر

(distribution free test) بمعــنى أنــه لايشــترط تحقق توزيع معين للظاهرة فى المجتمع.

(٢) أن هذا التوزيع – مع توافر الشروط الخاصة بعدد التكرارات المتوقعة فى كــل خلــية – يصــلح للاســتخدام فى حالة القياسات النوعية والترتيبية والقياســات النســبية أو بفــترة مــتى أمكن تصنيف أو تبويب القراءات المشاهدة فى فئات مختلفة.

(٣) إذا كــان جــدول التوافق الذى وزعت فيه مفردات العينات بحسب فئات المتغــيرات هو جدول ٢ × ٢ فيمكن تطبيق أسلوب كا^{٢} (٢ × ٢) الحالة الخاصة كما أوضحنا فى السابق.

د : استخدام توزيع كا^{٢} فى الاستنتاج الإحصائى بشأن تباين المجتمع $\sigma^{٢}$:

يستخدم توزيع كا^{٢} فى الاستنتاج الإحصائى بشأن تباين المجتمع $\sigma^{٢}$ وذلك اعتماداً على أن توزيع العينات للمتغير العشوائى.

$$\frac{(ن - ١)\ ع^{٢}}{\sigma^{٢}} \qquad (٧)$$

له توزيع كا^{٢} بدرجات حرية (ن – ١) متى كانت العينة العشوائية مسحوبة مــن مجــتمع معتاد (أو قريباً من المعتاد) وحيث ع^{٢} التباين محسوباً من العينة العشوائية بالعلاقة مجـ (س – $\overline{س}$)^{2} ÷ ن – ١ .

وتأسيساً على ذلك فإنه يمكن تقدير $\sigma^{٢}$ بفترة ثقة (١ – α) ١٠٠٪ كالآتى:

$\sigma^{٢}$ = (ن – ١) ع^{٢} ÷ كا$^{٢}_{\frac{\alpha}{٢}}$ كحد أدنى.

= (ن – ١) ع^{٢} ÷ كا$^{٢}_{١ - \frac{\alpha}{٢}}$ كحد أعلى.

كما يمكن استخدام المتغير العشوائى

$$كا^{٢} = \frac{(ن - ١)\ ع^{٢}}{\sigma^{٢}}$$

كأداة لاختبارات الفروض المتعلقة بـ σ^2 (أنظر تمرين ٨ ، ٩).

* **اختبار المتتابعات** The Runs Test :

نحـــن نفترض دائماً عند إجراء اختبار إحصائى ما أو عند قياس ظاهرة ما أن القياســـات مسجلة من مفردات عينة عشوائية وإن كن لا نتحقق من عشوائية العيـــنة وبالـــتالى فإنه إن لم تكن العينة عشوائية فإن المقاييس الإحصائية التى تحســـب قـــد تكـــون متحـــيزة كما أن القرارات التى تتخذ فى شأن الاختبارات الإحصائية قد تكون غير صحيحة.

فـــإذا ألقيـــت قطعة عملة ٢٠ مرة فى تجربة عشوائية لتسجيل عدد مرات ظهـــور الوجه الذى يحمل الصورة أو الكتابة وبفرض أن نتيجة التجربة كانت كالآتى:

ص ص ص ص ص ص ص ص ص ص
ك ك ك ك ك ك ك ك ك ك

أو كانت على النحو التالى:

ص ك ص ك ص ك ص ك ص ك
ص ك ص ك ص ك ص ك ص ك

فإن كلا من الناتجين يشير إلى أنهما نواتج غير عشوائية إذ أن تتابع ظهور الصـــورة والكتابة فى الحالتين لا يتفق مع العشوائية لــ ن = ٢٠ حيث تشمل السلسة الأولى على متتابعتين وهو عدد قليل قد يكون نادر التحقق وتشمل الثانية على ٢٠ متتابعة وهو عدد كبير لا يتحقق إلا باحتمال صغير للغاية.

فإذا عرفت المتتابعات (Runs) بأنها ظهور قيمة أو وجه من أوجه المتغير مـــرة أو أكثر مسبوقة أو متبوعة بظهور القيمة أو الوجه الآخر للمتغير مرة أو أكثر فإن توزيع العينات للمتتابعات يعطى مؤشراً يصلح للتعرف أو الحكم على عشوائية العينة وبالتالى أنها تمثل المجتمع الذى سحبت منه.

١: اختبار المتتابعات فى حالة العينات الصغيرة ن ≥ ٢٠ :

خطوات الاختبار:

دع ن١= عدد العناصر من نوع ما أو وجه من أوجه المتغير،

ن٢ = عدد العناصر من نوع ثان من المتغير أو وجهه الآخر ،

ن١ + ن٢ = ن = حجم العينة.

مــ : عدد المتتابعات.

فــإذا كانــت ن١ ، ن٢ كلاهمــا ≥ ٢٠ فإن ملحق رقم (٨) يعطى الحدين الأدنـــى والأعلـــى لقــبول العدد المشاهد من المتتابعات والحكم بأن العينة عشــوائية. وأى عــدد يقل عن الحد الأدنى أو يزيد عن الحد الأعلى يعنى أن احــتمال مشاهدة عدد من المتتابعات يساوى أو يقل عن الحد الأدنى = $\frac{١}{٢}\alpha$. كمــا أن احتمال مشاهدة عدد من المتتابعات مساو للحد الأعلى أو أكبر منه = $\frac{١}{٢}\alpha$. وفــى كلــتا الحالتيــن يرفض الفرض بعشوائية العينة عند مستوى المعنوية α مما يترتب عليه رفض الفرض بعشوائية العينة.

فمثلاً عند ن١ = ن٢ = ١٠ وعند α = ٥٪ فإن الحد الأدنـــى لــ مــ ≥ ٦ والحــد الأعلى ≤ ١٦ واحتمال الحصول على مــ ≥ ٦ أو مــ ≤ ١٦ كل يتحقق باحتمال قدره ٠,٠٢٥ . هذا ويمكن إجراء الاختبار فى طرف واحد.

مثال (٨):

للتعرف على رأى قائدى السيارات فى إجراءات تنظيم المرور وأثره على إنسياب حركة المرور فقد سجل رأى عشرون من قائدى السيارات عند تقاطعين رئيســيين فــى مديــنة القاهرة وكانت إجابات أفراد العينة بالموافقة (نعم) أو الرفض (لا) كما هو موضح بجدول (٨) التالى:

جدول (٨)

توزيع أراء ٢٠ من قائدى السيارات بشأن تنظيم المرور

مسلسل	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الإجابة	نعم	نعم	لا	لا	لا	نعم	لا	لا	نعم	نعم
المتتابعات	٦			٢	.	٣	٤		٥	
مسلسل	١١	١٢	١٣	١٤	١٥	١٦	١٧	١٨	١٩	٢٠
الإجابة	لا	لا	لا	نعم	لا	نعم	لا	لا	لا	نعم
المتتابعات		٦		٧	٨	٩		١٠		١١

هــل هناك شك فى أن هذه الإجابات صدرت عن عينة عشوائية من قائدى السيارات عند α = ٥٪ ؟.

الحل :

ن$_{١}$ = ٨ ، ن$_{٢}$ = ١٢

مـ* = ١١ متتابعة

الفرض الأصلى : أن العينة من الإجابات العشوائية.

عــند α = ٥٪ فإن مـ* ≥ ٦ ، مـ* ≤ ١٦ تؤدى إلى رفض الفرض بعشوائية العينة.

ولكــن مـ* = ١١ لذلــك يقــبل الفرض عند α = ٥٪ وهذه العينة من الإجابات تؤكد عشوائية العينة عند مستوى المعنوية ٥٪ .

ب : اختبار المتتابعات فى حالة العينات الكبيرة :

كان الاختبار السابق فى حالة العينات الصغيرة ، أى حين تكون كــل من ن$_{١}$ ، ن$_{٢}$ < ٢٠ أما فى حالة العينات الكبيرة ، أى حين تزيد قيمة كل من ن$_{١}$ ، ن$_{٢}$ عــن ٢٠ فإنه يمكن استخدام التقارب الإعتدالى فى اختبار العشوائية إذ أن توزيع العينات للمتتابعات توزيع معتاد:

$$\text{توقعه } \mu_{\text{ـ}} = \frac{\text{٢ن}_{١} \times \text{ن}_{٢}}{\text{ن}_{١} + \text{ن}_{٢}} + ١ \qquad (٨)$$

$$\text{وإنحرافه المعيارى } \sigma_{\text{ـ}} = \sqrt{\frac{\text{٢ن}_{١}\text{ن}_{٢}(\text{٢ن}_{١}\text{ن}_{٢} - \text{ن}_{١} - \text{ن}_{٢})}{(\text{ن}_{١} + \text{ن}_{٢})^{٢}(\text{ن}_{١} + \text{ن}_{٢} - ١)}} \qquad (٩)$$

وبالتالى فإن المتغير العشوائى : $\frac{\text{ـ} - \mu_{\text{ـ}}}{\sigma_{\text{ـ}}}$ ←توزيع ى: م (صفر ،١) (١٠)

وبالتالى يمكن استخدام التوزيع المعتاد المعيارى فى اختبار العشوائية.

مثال (٩):

أعلنــت مؤسسة ما عن طلب عاملين من الذكور والإناث للعمل بالعلاقات العامة وكان ترتيب ورود الطلبات المقدمة للمؤسسة بحسب النوع (ذكور : ذ ، إناث : أ) كالآتى :

ذ أ ذ أ ذ ذ ذ أ أ ذ أ ذ أ
ذ أ ذ ذ ذ ذ أ ذ أ ذ أ ذ ذ
أ أ أ ذ أ ذ أ ذ أ ذ ذ أ ذ
ذ أ ذ ذ ذ ذ أ ذ أ ذ ذ

إختبر هذه السلسلة من حيث العشوائية عند α = ١٪ .

الحل:

ن$_{١}$ = ٣٠ ذكور ، ن$_{٢}$ = ٢٠ إناث

الفرض الأصلى : أن الطلبات عشوائية فى تسلسل ورودها عند α = ١٪

ويطــبق اختــبار المتتابعات للعينات الكبيرة ويرفض الفرض الأصلى إذا كانت :

$$\text{س} = \frac{\text{ـ} - \mu_{\text{ـ}}}{\sigma_{\text{ـ}}} < -٢,٥٨ \text{ أو } > ٢,٥٨$$

ومن بيانات العينة

$$م^{*} = ٣٥$$

$$\mu_{م} = \frac{٢ن_{١}\ ن_{٢}}{ن_{١} + ن_{٢}} + ١ = \frac{٢\times٣٠\times٢٠}{٣٠ + ٢٠} + ١ = \frac{١٢٠٠}{٥٠} + ١ = ٢٥$$

$$\sigma_{م} = \frac{٢ن_{١}\ ن_{٢}\ (٢ن_{١}\ ن_{٢} - ن_{١} - ن_{٢})}{(ن_{١} + ن_{٢})^{٢}\ (ن_{١} + ن_{٢} - ١)}$$

$$= \frac{٢\times٣٠\times٢٠\ (١٢٠٠ - ٥٠)}{٢٥٠٠\times٤٩} = ١١,٢٦٥$$

$$\sigma_{م} = ٣,٣٥٦$$

$$ى^{*} = \frac{٣٥ - ٢٥}{٣,٣٥٦} = ٢,٩٨ > ٢,٥٨$$

ويـــرفض الفــرض بعشوائية الطلبات المقدمة بحسب النوع وذلك بحسب تسلسل ورودها عند $\alpha = ١\%$

ملحوظـة:

ويلاحـــظ انـــه بسبب طبيعة القياسات وأنها ترتيبية (بحسب تتابعها) فإنه لا يوجد اختبار معلمى أو بارامترى يصلح للاستخدام كبديل لاختبار المتتابعات لذلك فليس هناك مبرر عن قوة اختبار المتتابعات.

(٣) إختبار الإشارة Sign Test:

١- الاختبار فى حالة عينة واحدة:

يســـتخدم اختبار الإشارة لعينة واحدة لاختبار الفروض الإحصائية المتعلقة بمقــياس للنزعة المركزية (Central Tendency) فى المجتمع سواء كان هذا المقــياس هو الوسط الحسابى μ أو الوسيط ولذلك لم نقل بأنه اختبار لابارامترى بشــأن الوسط الحسابى. ويفترض هنا أن العينة سحبت عشوائياً من مجتمع تتبع فيه الظاهرة موضوع الاختبار توزيعاً مستمراً ولكننا لا نفترض فرضاً ما بشأن

شـكل الـتوزيع ولذلـك فهو اختبار بديل لاختبارات الخاص بالوسط الحسابى للمجتمع μ الذى يفترض أن يكون التوزيع فى مجتمع العينة معتاداً.

وحيـن يكـون الـتوزيع فى مجتمع الدراسة مستمراً ومتماثلاً أو قريباً من الـتماثل بحيث يكون احتمال أن تأخذ أحد مفردات العينة العشوائية قيمة ما تقل عن الوسط الحسابى μ مساو لاحتمال أن تأخذ قيمة أكبر من μ وهذا الاحتمال $= \frac{١}{٢}$ وفى هذه الحالة فإن مقياس النزعة المركزية موضوع الاختبار باستخدام اختـبار الإشارة هو الوسط الحسابى فى المجتمع μ أو الوسيط فى المجتمع أيهما يصـلح لهـذا الفرض. أما إذا كان التوزيع ملتوياً فإن مقياس النزعة المركزية موضـوع الاختبار سيكون الوسيط فى المجتمع اعتماداً على أن تعريف الوسيط بأنـه القيمة التى تتوسط التوزيع وأن يكون عدد المفردات الأكبر منها فى القيمة مساو لعدد المفردات الأصغر فى القيمة.

هذا بالإضافة إلى عدم تأثر الوسيط بالقيم القليلة المتطرفة كالوسط الحسابى. وأمـا إذا كان التوزيع متماثلاً أو قريباً من التماثل فإن الوسط الحسابى والوسيط يكونا متساويين فى القيمة أو يقتربان من التساوى ولذلك فيصلح أيهما للإستخدام مع اختبار الإشارة.

<u>خطوات الاختبار:</u>

(١) الفرض العدمى هو $\mu = \mu_{٠}$ أو أن الوسيط ($ر_{٢}$) = قيمة معينة مقابل فرض بديل مناسب.

(٢) يحدد مستوى المعنوية المناسبة α .

(٣) ثم تستبدل قيم العينة التى تقل عن $\mu_{٠}$ بالإشارة (−) والقيم التى تزيد عن $\mu_{٠}$. (أو الوسيط $ر_{٢}$) بالإشارة (+). وبذلك يكون الاختبار هو اختبار لمتغير له توزيع ذو الحدين حيث $ل = \frac{١}{٢}$. وتسـتبعد المفردات التى تتساوى فى القيمة مع $\mu_{٠}$.

أ: اختبار العينة فى حالة العينات الصغيرة:

إذا كان حجم العينة نَ $\leq$ ٢٠ اعتبرت العينة صغيرة وبالتالى فيمكن تحديد إحــتمال مشاهدة الإشارات (+) أو (-) بعدد مساو للعدد المشاهد أو أكبر منه أو أقل منه ويقبل الفرض العدمى أو يرفض حسب الأحوال ويمكن استخدام اختبار الإشارة اعتماداً على توزيع ذو الحدين كما يوضحه المثال التالى:

مثال (١٠):

البــيانات التالــية تمثل متوسط عدد أيام الأجازات التى حصل عليها عمال المبــيعات فــى عيــنة عشوائية مكونة من ١٥ فرعاً من أفرع أحد المؤسسات التجارية وذلك خلال أحد السنوات:

٣٣,٩ ٣٥,٤ ٣٧,٣ ٤٠,٩ ٢٧,٨ ٣٥,٥ ٣٤,٦

٤١,١ ٣٠,٠ ٤٣,٢ ٣٣,٩ ٤١,٣ ٣٢,٠ ٣٧,٧ ٣٥,٢

اختبر الفرض $\mu_{.}$ = ٣٢ يوماً عند α = ٠,٠١

ســوف يحل المثال باستخدام اختبار الإشارة لعينة واحدة ثم بفرض أن كافة شروط استخدام اختبار ت قائمة فسوف نعيد المثال باستخدام اختبار ت:

الحل:

(١) اختبار الإشارة:

باستبدال القيم بالإشارات (+) لتلك التى تزيد فى قيمتها عن $\mu_{.}$ = ٣٢ ، (-) لــتلك التى تقل عن هذه القيمة واستبعاد المفردة الثالثة عشرة وتساوى $\mu_{.}$ = ٣٢ تكون الإشارات:

+ + + + - + + + - + + + + + وعددها ١٤ إشارة.

١٢(+) ، ١٢(-)

وإذا كان الفرض العدمى $\mu_{.}$ = ٣٢

والبديل μ > ٣٢

ويتوافر فى حالتنا هذه كافة الشروط الخاصة بالتوزيع ذو الحدين: وهى أن كل محاولة ينتج عنها أحد ناتجين إما نجاح (عدد أيام الأجازات > ٣٢ يوماً) أو فشـــل وهـــو العكس ، وكل باحتمال ثابت (ل = ٠,٥) وعدد المحاولات ن ثابت وهـــو ١٥ فـــى مثالنا هذا ، المحاولات المستقلة لذلك فإنه يمكن حساب احتمال الحصول على ١٢(+) أو ١٣(+) أو ١٤(+) وهو ما يعادل الحصول على ٢(-) أو اقل. وباستخدام ذو الحدين:

$$\text{ح}(\text{س} \leq ٢) = \sum_{\text{س}=\text{صفر}}^{٢} \binom{١٤}{\text{س}} \left(\frac{١}{٢}\right)^{\text{س}} \left(\frac{١}{٢}\right)^{٠,٠٢٤}$$

$$= \binom{١٤}{٠} \left(\frac{١}{٢}\right)^{٠} \left(\frac{١}{٢}\right)^{١٤} + \binom{١٤}{١} \left(\frac{١}{٢}\right)^{١} \left(\frac{١}{٢}\right)^{١٣} + \binom{١٤}{٢} \left(\frac{١}{٢}\right)^{٢} \left(\frac{١}{٢}\right)^{١٢}$$

ويمكـــن الحصـــول على قيمة هذا الاحتمال من ملحق رقم (٩) الذى يعطى احتمال الحصول على س أو أقل إشارة (الإشارة الأقل عدداً وهى (-) فى مثالنا) حيث $\text{ل} = \text{م} = \frac{١}{٢}$ وإلا استخدمت الصيغة.

$$\text{مجـ} \binom{\text{ن}}{\text{س}} \text{ل}^{\text{س}} \text{م}^{\text{ن}-\text{س}} \qquad (١١)$$

للحصول على الاحتمال المطلوب متى كانت $\text{ل} \neq \text{م} \neq \frac{١}{٢}$

وفى مثالنا هذا:

ح (س = ٠ ، ١ ، ٢ | ن = ١٤) = ٠,٠٠٦ + ٠,٠٠١

= ح (+ = ١٢ إشارة أو أكثر)

= ٠,٠٠٧ < ٠,٠١ لذلك يرفض الفرض العدمى.

ملحوظة:

إذا كان الفرض العدمى μ. = ٣٢ يوماً.

والفرض البديل $\mu \neq$ ٣٢ ، < ٣٢ أو > ٣٢ أى اختبار طرفين.

فإنـــه يمكـــن تكـــرار نفـــس خطوات الاختبار السابق مع مضاعفة احتمال الحصول على:

(س = ٠ ، ١ ، ٢ | ن = ١٤) = ٢ × ٠,٠٠٧ = ٠,٠١٤

٢: اختبار ت:

س̄ = ٥٣٩,٨ / ١٥ = ٣٥,٩٨٧ يوماً

ع = ٥٠٤,١١ / √١٤ = ٦,٠٠١

الفرض العدمي μ_0 = ٣٢ يوماً   الفرض البديل μ > ٣٢

α=٠,٠١ وعند α = ١٪ يرفض الفرض العدمي إذا كانت ت* ≤ ٢,٦٢٤

ومن بيانات العينة:

ت* = (س̄ − μ_0) / (ع / √ن) = (٣٥,٩٨٧ − ٣٢) / (٦,٠٠١ / √١٥)

= ٢,٥٧٤

لذلك يقبل الفرض العدمي عند α = ١٪ .

واضح من هذا أن القرار قد اختلف في حالة استخدام الإشارة في اختبار ت حيث كان إجراء اختبار الإشارة (أو ما يشار إليه أحياناً باسم اختبار ذو الحدين The Binomial Test) على أساس استبدال القيم بإشارت موجبة وسالبة الذى حول التوزيع إلى توزيع ذو الحدين.

ومع ذلك فإنه ما لم تكن القراءات وصفية أو نوعية مما لا يصلح معها استخدام اختبار ت فإن اختبار الإشارة يكون أقل قوة إذ تصل قوته المكافأة إلى ٩٥٪ وتتناقص حتى تصل إلى ٦٣٪ . وإذا كانت القراءات يصلح معها تطبيق اختبار ت فإنه يمكن اختبار مطابقة توزيعها للتوزيع المعتاد – باستخدام اختبار كا^{٢} كما أوضحنا في بداية هذا الفصل – وإذا تبين أن توزيعها يتبع التوزيع المعتاد أو لا يختلف كثيراً عنه فإنه يمكن استخدام اختبار ت وإلا استخدم اختبار الإشارة.

٢: اختبار الإشارة فى حالة العينات الكبيرة:

إذا كانـــت ن > ٢٠ اعتـــبرت العيـــنة كبيرة وبدلاً من استخدام التوزيع ذو الحديـــن لإجراء الاختبار فإنه يمكن الاعتماد على التقارب الإعتدالى للتوزيع ذو الحديـــن مـــتى توافر الشرط اللازم وهو أن ن م أو ن ل كلاهما أكبر من ٥ فى هذه الحالة فإن أداة الاختبار هى:

$$\text{ى} \sim \frac{\text{س} - \text{ن ل.}}{\sqrt{\text{ن ل. م.}}} = \frac{(\text{س} \pm ٠,٥) - \text{ن ل.}}{\sqrt{\text{ن ل. م.}}} \qquad (١٢)$$

وقد استخدمت العلامة ~ للدلالة على " تؤول إلى ". أما ن ل ، ن ل م فهى توقع وتباين س وهو متغير عشوائى له توزيع ذو الحدين - وليس نسبة - أما ل . ، م. المعرفتين بفرض العدم.

ويمكن استخدام أى من الصيغتين السابقتين الأولى بدون تقريب أو تصحيح للإســتمرارية والثانية مصححة حيث س + ٠,٥ إذا كانت س < ن ل. ، س - ٠,٥ إذا كانت س > ن ل. .

مثال (١١)[*] :

البـــيانات التالية توضح كمية العادم اليومى من أكسيد الكبريتيك الذى يخرج من أحد الوحدات الصناعية الكبيرة (بالطن):

١٧	١٥	٢٠	٢٩	١٩	١٨	٢٢	٢٥	٢٧	٩
٢٤	٢٠	١٧	٦	٢٤	١٤	١٥	٢٣	٢٤	٢٦
١٩	٢٣	٢٨	١٩	١٦	٢٢	٢٤	١٧	٢٠	١٣
١٩	١٠	٢٣	١٨	٣١	١٣	٢٠	١٧	٢٤	١٤

استخدم اختبار الإشارة لعينة واحدة لاختبار الفرض بأن μ_0 = ٢٣,٥ طن مقابل الفرض بأن μ < ٢٣,٥ عند α = ٥٪ .

(*) Freund, J. E. & William, F.J. : Elementary Business Statistics : The Modern Approach, Prentice-Hall International Ed., 4th ed., p. 411., 1964 .

الحـل:

عدد الإشارات + = ١١

عدد الإشارات − = ٢٩

ولأن ن ل = ن م = ٤٠ × = ٢٠ > ٥

لذلك يمكن استخدام اختبار الإشارة (ذو الحدين) مع التقارب الإعتدالى.

$$\text{ى}_{٠} = \frac{\text{س} - \text{ن ل}}{\sqrt{\text{ن ل م}}} = \frac{١١ - ٤٠ \times \frac{١}{٢}}{\sqrt{٤ \times \frac{١}{٢} \times \frac{١}{٢}}} = -٢,٨٥$$

وهى أصغر من −١,٦٤٥ لذلك يرفض الفرض العدمى.

ب: اختبار الإشارة فى حالة عينتين غير مستقلتين:

ويناظر هذا الأسلوب اختبار القراءات المزدوجة (Paired Observations) أى حين تسجل القراءتين نفس المفردة. ويكون الهدف فى حالتنا هذه هو اختبار مؤثر ما وتكون القراءة الأولى على مفردة العينة قبل استخدام أو إدخال المؤثر الذى يراد قياس أثره ثم تسجل القراءة الثانية على نفس المفردة بعد إدخال المؤثر ويستخدم الفرق بين القراءة الأولى والثانية على مفردات العينة فى اختبار معنوية أثر المؤثر موضوع الاختبار أو القياس.

وبديهى أن القراءات المزدوجة غير مستقلة لتسجيلها على نفس المفردات. ويمكن استخدام أسلوب بارامترى لاختبار الفرق بين المتوسطين ويعتمد هذا الأسلوب على اختبار ت للفروق بين أزواج القيم كما أوضحنا من قبل ، أما المدخل الآخر فهو اختبار لا بارامترى ويعتمد على اختبار الإشارة وسوف نعالج هذين المدخلين بالمثال التالى:

مثال (١٣)(*) :

لاختـــبار مدى فعالية نظام جديد لضبط المرور قد سجل عدد الحوادث التى رصدت عند ثمانية تقاطعات رئيسية اختيرت عشوائياً ، وذلك خلال أربعة أسابيع قبل وبعد استخدام النظام الجديد. وكانت النتائج كالتالى:

قبل	٩	٧	٣	١٦	١٢	١٢	٥	٦
بعد	٥	٣	٤	١١	٧	٥	٥	١

اختـــبر عند مستوى المعنوية ١٠٪ أن النظام الجديد للمرور لا يختلف عن الـــنظام السابق مقابل الفرض البديل أن أعلى كفاءة. والاختبار باستخدام اختبار الإشارة كالآتى:

الحـل:

(١) الفرض $\mu_1 = \mu_2$ مقابل $\mu_1 > \mu_2$

(٢) α = ٠,١

(٣) باســـتخدام اختـــبار الإشارة ، وتكون الإشارة +١ إذا كانت القراءة الأولى أكبر من الثانية ، – فى حالة العكس.

ويـــرفض الفـــرض العدمـــى إذا كان احتمال الحصول على عدد + ≥ ٦ مرات < ٠,١ .

(*) Freund. J. E. & William. F.J. & B.M.. :
Elementary Business Statistics : The Modern Approach.. 6th ed..
Prentice-Hall. p. 568. 1994 .

جدول (٩)

عدد حوادث المرور المسجلة عند ٨ تقاطعات رئيسية

بحسب نظام المرور

قبل	بعد	الإشارة	الفرق (ف)	$ف^2$
٩	٥	+	٤	١٦
٧	٣	+	٤	١٦
٣	٤	−	١−	١
١٦	١١	+	٥	٢٥
١٢	٧	+	٥	٢٥
١٢	٥	+	٧	٤٩
٥	٥	تستبعد	صفر	صفر
٦	١	+	٥	٢٥
			٢٩	١٥٧

عدد الإشارات + = ٦

ومن جدول ملحق رقم (٩)

ح (س + ≤ ٦) = ح (س − ≥ ١) = ٠,٠٦٣ > ٠,١

لذلك يرفض الفرض العدمى أى أن النظام الجديد أعلى كفاءة فى ضبط عدد حوادث المرور عن النظام القديم.

وباستخدام اختبار ت للقرارات المزدوجة:

$$\text{ت}^* = \frac{\bar{\text{ف}} - \text{صفر}}{\text{ع ف}} = \frac{\text{٣,٦٢٥} - \text{صفر}}{\sqrt{\frac{\text{٧,٤١٠٧}}{\text{٨}}}} = \frac{\text{٣,٦٢٥}}{\text{٠,٩٢٦٣}} = \text{٣,٩١٢}$$

ولكن $\text{ت}_{(\text{٠,٩٠})}$ = ١,٤١٥ > ت^* = ٣,٩١٢

لذلــك يرفض الفرض بأن $\mu_1 = \mu_2$ وتتفق نتيجة هذا الاختبار مع نتيجة اختــبار الإشارة ، ولكن الاختبار البارامترى أعلى قوة من اختبار الإشارة لذلك يفضل الاختبار الأخير متى توافرت شروط صحة استخدامه.

(٤) اختبار مجموع الرتب / اختبار U: مان-ويتنى

The Mann-Whitney U Test

سبق أن قدمنا اختبارات بارامترية لاختبار الفرق بين متوسطين $\mu_1 = \mu_2$. اعتماداً على وسطين حسابيين $\bar{س}_1$ ، $\bar{س}_2$ حسبا من عينتين مستقلتين سواء كان التباين للمجتمعين معلوماً أو غير معلوم وفى الفقرة التالية نقدم اختباراً لا معلمياً (لا بارامترى) يعتمد على الرتب لاختبار الفرق بين المتوسطين لعينتين مستقلتين سحبتا كل من مجتمع له توزيع مستمر وليس بالضرورة معتاداً.

وهــناك مجموعــة الاختبارات على مجموع الرتب المندمجة للعينتين وذلك لاختبار – ما إذا كانت عينتين عشوائيتين مستقلتين قد سحبتا من مجتمع واحد أم لا هــذا متى كانت القياسات ترتيبية أو يمكن استخدامها فى ترتيب القراءات. بل ويمــيل الباحثون إلى تفضيل استخدام هذا الاختبار بدلاً من اختبارات للفرق بين المتوسطين ($\mu_1 = \mu_2$) متى كانت شروط استخدام اختبارات غير مؤكدة.

خطوات الاختبار:

(١) تدمـــج مفــردات العينتين المستقلتين معاً (مع تمييزهما بحسب العينة التى تنــتمى إليها كل) ثم تستبدل المفردات أو المشاهدات بالرتب مرتبة ترتيباً تصـــاعدياً أو تنازلــياً – مع إعطاء المفردات المكررة المتوسط الحسابى للرتب المناظرة.

(٢) وإذا كانــت $ن_1$ ، $ر_1$ ترمــز إلـــى عــدد مفردات ومجموع رتب العينة الأصغر حجماً ، $ن_2$ ، $ر_2$ إلى عدد مفردات ومجموع رتب العينة الأكبر حجماً فإن أداة الاختبار تتحدد على أساس قيمة $ن_2$.

أ: اختبار U فى حالة العينات الصغيرة:

فإذا كانت $ن_٢ \leq ٢٠$ استخدم مدخل " العينات الصغيرة " فى اختبار الفرض بأن العينتين تنتميان إلى مجتمع واحد ، أى أن التوزيع واحد للظاهرة فى العينتين. وتستخدم أداة الاختبار.

$$U_١ = ن_١ ن_٢ + \frac{ن_١ (ن_١ + ١)}{٢} - ر_١ \qquad (١٣)$$

أو

$$U_٢ = ن_١ ن_٢ + \frac{ن_٢ (ن_٢ + ١)}{٢} - ر_٢ \qquad (١٤)$$

وتقارن القيمة المحسوبة لـ U بالقيمة المعنوية لـ U_α كما يحددها جدول ملحق رقم (١٠) ويرفض الفرض العدمى إذا كانت $U^* \leq U_\alpha$ ويقبل فيما عدا ذلك من الأحوال. وعادة تستخدم U الأصغر قيمة فى الاختبار. وإذا تبين أن القيمة المحسوبة $U^* > \frac{ن_١ ن_٢}{٢}$ فإنها ستكون الأكبر قيمة ولإيجاد القيمة الأصغر لـ U تستخدم العلاقة التالية:

$$U = ن_١ ن_٢ - U^* \qquad (١٥)$$

فمثلاً إذا حسبت $U_١$ وتبين أنها $> \frac{ن_١ ن_٢}{٢}$ فبدلاً من حساب $U_٢$ من (١٤) فإنه يمكن إيجاد قيمتها من العلاقة (١٥) أى أن :

$$U_٢ = ن_١ ن_٢ - U_١$$

ب: اختبار U فى حالة العينات الكبيرة:

أما إذا كانت $ن_٢ \leq ٢٠$ فيستخدم المدخل الخاص بالعينات الكبيرة الذى يعتمد على التقارب الاعتدالى لتوزيع المتغير العشوائى U إذ أن توزيع العينات للمتغير العشوائى U يقترب من التوزيع المعتاد الذى متوسطه

$$\mu_U = \frac{ن_١ ن_٢}{٢} \qquad (١٦)$$

وتباينه σ^{2}_{u} = $\frac{(ن_{١} ن_{٢}) (ن_{١}+ن_{٢}+١)}{١٢}$ (١٧)

وبالتالى فإن ى = $\frac{U - \mu_{U}}{\sigma_{U}}$ له توزيع يؤول إلى التوزيع

المعتاد المعيارى: م (صفر ، ١) (١٨)

وعلـيه فتسـتخدم أداة الاختـبار (١٨) فى اختبار الفرض فى هذه الحالة. والأمـثلة التالـية توضح كيفية استخدام هذا الاختبار فى حالة العينات الصغيرة والكبيرة.

مثال (١٣):

البـيانات التالـية تبين الطاقة الحرارية (مليون سعر حرارى / للطن) لعينة عشوائية من الفحم من منجمين مختلفين:

المنجم أ:	٨٣٨٠	٨٢١٠	٧٣٦٠	٧٨٤٠	٧٩١٠
المنجم ب:	٧٥٤٠	٧٧٢٠	٧٧٥٠	٨١٠٠	٧٦٩٠

اختبر الفرض بأن العينتين من مجمعين متماثلين عند α = ٥٪ .

الحـل:

الفــرض العدمى: أن العينتين من منجمين متماثلين تماماً أما الفرض البديل فهو أن المنجمين يختلفان فى متوسط الطاقة الحرارية للفحم المستخرج.

α = ٥٪

وبترتيب المفردات فى العينتين معاً ترتيباً تصاعدياً نجد أن:

	٧٥٤٠	٧٦٩٠	٧٧٦٠	٧٧٥٠	٧٨٤٠	٧٩١٠	٨١٠٠	٨٢١٠	٨٣٦٠	٨٣٨٠
منجم	ب	ب	ب	ب	أ	أ	ب	أ	أ	أ
ترتيب	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠

$ر_{أ}$ = ٣٨ $ر_{ب}$ = ١٧

$$U^*_{\text{ب}} = \text{ن}_{١}\,\text{ن}_{٢} + \frac{\text{ن}_{٢}\,(\text{ن}_{٢}+١)}{٢} - \text{ر}_{\text{ب}}$$

$$= ٥ \times ٥ + \frac{٥\times٦}{٢} - ١٧ = ٢٣ > \frac{\text{ن}_{١}\,\text{ن}_{٢}}{٢} = \frac{٢٥}{٢} = ١٢,٥$$

وبالتالى فإن $U^*_{\text{ب}}$ هى الأكبر قيمة ولحساب قيمة $U^*_{\text{أ}}$ (أى الأصغر قيمة) باستخدام العلاقة (١٥).

$$U^*_{\text{أ}} = ٢٥ - ٢٣ = ٢$$

وهـــى = $U_{٥\%}$ طرفيـــن كمـــا هو مبيـــن بجدول ملحـــق رقم (١٠) لذلك يرفض الفرض بعدم وجود اختلاف فى التوزيع فى المجتمعين.

وجدير بالذكر أنه باستخدام (١٣) فإن:

$U^*_{\text{أ}} = ٥ \times ٥ + \frac{٥\times٦}{٢} - ٣٨ = ٢$ وهى نفس القيمة السابق حسابها لـــ $U^*_{\text{أ}}$

ولو طبقنا اختبار ت على بيانات نفس المثال حيث:

الفرض العدمى $\mu_{١} = \mu_{٢}$ والفرض البديل $\mu_{\text{أ}} \neq \mu_{\text{ب}}$. وعند $\alpha = ٥\%$ فإن:

$$\text{ت}^* = ٢,٦١ < \text{ت}_{٨،٠,٠٥} = ٣,٣٥٥$$

لذلـــك يقبل الفرض العدمى بعدم اختلاف الطاقة الحرارية المتوسطة للفحم المســـتخرج من المنجمين. وفى الحقيقة فإن التباين فى العينة الأولى حوالى مرة ونصف التباين فى العينة الثانية مما يعيب استخدام اختبار ت.

مثال (١٤):

ســـحبت عينتيـــن عشـــوائيتين الأولى من عمال قسم المبيعات فى صناعة للكيماويات والأخرى عمال قسم بيع أدوات منزلية وسجلت الأجور اليومية لأفراد كل من العينتين فكانت النتائج كالآتى:

جدول (١٠)

الأجور اليومية لعينتين كل من عمال المبيعات فى قسمين مختلفين

العينة الأولى:

٧,٦٠٠	١٠,٤٠٠	١٢,٥٠٠	٧,٣٠٠	٨,٧٠٠	٦,٩٠٠
٩,٧٠٠	١٤,٩٠٠	٧,٨٠٠			

العينة الثانية:

٨,٨٠٠	٩,٢٠٠	١٠,١٠٠	٨,٣٠٠	٩,٩٠٠	٨,٥٠٠
١١,١٠٠	٧,١٠٠	٩,٠٠٠	٩,٤٠٠	٩,١٠٠	

اختبر الفرض بعدم وجود فرق بين الأجر اليومى لعمال القسمين.

الحـل:

تدمج بيانات العينتين معاً ثم ترتب المفردات ترتيباً تصاعدياً (أو تنازلياً) مع إعطاء متوسط الرتب للمفردات المكررة فمثلاً:

جدول (١١)

الترتيب التصاعدى لبيانات جدول (١٠) للعينتين

٦,٩٠٠	٧,١٠٠	٧,٣٠٠	٧,٦٠٠	٧,٨٠٠	٨,٣٠٠
١	٢	١	١	١	٢
٨,٥٠٠	٨,٧٠٠	٨,٨٠٠	٩,٠٠٠	٩,١٠٠	٩,٢٠٠
٢	١	٢	٢	٢	٢
٩,٤٠٠	٩,٧٠٠	٩,٩٠٠	١٠,١٠٠	١٠,٤٠٠	١١,١٠٠
٢	١	٢		١	٢
١٢,٥٠٠	١٤,٩٠٠				
١	١				

(استخدم الرقم ١ للإشارة إلى العينة الأولى والرقم ٢ إلى العينة الثانية)

وعلى ذلك فإن مفردات كل من العينتين تحتل الرتب التالية

الأولى	١	٣	٤	٥	٨	١٤	١٧	١٩	٢٠		
الثانية	٢	٦	٧	٩	١٠	١١	١٢	١٣	١٥	١٦	١٨

وإذا صــح فرض عدم وجود فرق بين العينتين فإنه يتوقع أن يكون مجموع رتب مفردات كل من العينتين متقارباً أو متساوياً.

ولأن $ن_٢$ = ١١ < ٢٠ لذلك يستخدم أسلوب اختبار العينة الصغيرة.

$$U_1^* = ٩ \times ١١ + \frac{٩ \times ١٠}{٢} - ٩١ = ٥٣ > \frac{ن_١ ن_٢}{٢} = ٤٩,٥$$

وهى الأكبر قيمة ولإيجاد U الأصغر قيمة وهى $U_٢$

$$U_٢^* = ن_١ ن_٢ - U_1 = ٩٩ - ٥٣ = ٤٦$$

(وللتحقق $U_٢ = ٩ + \frac{١٣٢}{٢} - ١١٩ = ٤٦$)

وبمقارنة $U_٢^* = ٤٦$ بـــ $U_{٥٪}$ طرفين = ٢٣ أى أن $U^* > U_\alpha$

لذلك يقبل الفرض بعدم وجود اختلاف معنوى بين الأجور اليومية للعمال فى القسمين بمعنى أنهما ينتميان إلى مجتمع واحد أو مجتمعين متساويين.

مثال (١٥):

الجــدول الــتالى (١٢) يبين ترتيب مفردات عينتين عشوائيتين من الذكور والإناث فى اختبار للقدرات الميكانيكية.

جدول (١٢)

الذكـــــــور			الإنــــــــاث				
٢٦	٦	٣٨	٧	٣٣	٣٩	١٦	١٩
١٠	١٤	٢٤	٢٠	١	٢	٣٧	٢٩
٣٠	١٧	٤٠	٢٧	٩	٥	٢٣	٣١
٣	٢٢	٢٥	٢٨	١٣	٣٦	١٥	١٨
٣٢	٣٤	١٢	٤	٢١	٨	١١	٣٥

هــل تبين هذه الرتب اختلافات معنوية فى القدرات الميكانيكية بين الذكور والإناث عند مستوى المعنوية ٥٪ ؟

الحـل:

$ن_1$: للذكور = ١٥ $\qquad$ $ن_2$: للإناث = ٢٥

$ن_2$ > ٢٠ يستخدم التقارب الإعتدالى لــ U فى الاختبار.

$$\mu_U = \frac{\text{ن}_1 \text{ن}_2}{٢} = \frac{١٥ \times ٢٥}{٢} = ١٨٧,٥$$

$$\sigma_U = \frac{(\text{ن}_1 \text{ن}_2)(\text{ن}_1+\text{ن}_2+١)}{١٢}$$

$$= \frac{١٥ \times ٢٥ \; (٤١)}{١٢} = ١٢٨١,٢٥$$

$$\sigma_U = ٣٥,٧٩٥$$

$$U_1^* = \text{ن}_1 \text{ن}_2 + \frac{\text{ن}_1 (\text{ن}_1+١)}{٢} - \text{ر}_1$$

$$= ١٥ \times ٢٥ + \frac{١٥ \times ١٦}{٢} - ٣٣٣$$

$$= ٣٧٥ + ١٢٠ - ٣٣٣ = ١٦٢$$

$$\therefore \text{ى}^* = \frac{١٦٢ - ١٨٧,٥}{٣٥,٧٩٥} = -٠,٧١٢ \qquad -١,٩٦ < -٠,٧١٢ < ١,٩٦$$

لذلــك يقــبل الفــرض بعدم وجود اختلاف معنوى بين القدرات الميكانيكية للذكور والإناث عند α = ٥٪ .

وجدير بالذكر أن استخدام U_2 بدلاً من U_1 فى الاختبار سوف لا يؤثر إلا علـــى إشـــارة ى* فتتغـــير من − إلى + دون أن يؤثر ذلك على القرار بشأن الفرض موضوع الاختبار ويتأكد ذلك من الآتى:

$$U_2^* = \text{ن}_1 \text{ن}_2 + \frac{\text{ن}_2 (\text{ن}_2+١)}{٢} - \text{ر}_2$$

= ٣٧٥ + ٣٢٥ - ٤٨٧ = ٢١٣

= ن١ ن٢ - U_1^* = ٣٧٥ - ١٦٢ = ٢١٣

∴ ى* = $\frac{\text{٢١٣}-\text{١٨٧,٥}}{\text{٣٥,٧٩٥}}$ = ٠,٧١٢ ∴ -١,٩٦ < ٠,٧١٢ < ١,٩٦

ويقبلِ الفرض العدمى عند α = ٥٪ .

وهذا يؤكد أيضاً الإتجاه إلى U^* الأقل قيمة فى الاختبار.

جـ: ملاحظات على اختبار U:

(١) يفترض لجواز استخدام هذا الاختبار - كبديل للاختبار البارامترى ت للفرق بين متوسطين - أن يكون توزيع المتغير فى المجتمعين مستمراً وليس بالضرورى أن يكون معتاداً. وأن تكون العينتين مستقلتين.

(٢) ويفضل استخدام هذا الاختبار بدلاً من اختبار ت إذا كان واضحاً من البداية أن العينتين مختلفتى التباين بشكل واضح مما لا يصلح معه استخدام اختبار ت الذى يفترض أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ وبالتالى يمكن تقدير σ^2 بالتباين التجميعى $ع_ت^2$ من بيانات العينتين. كما يفضل أيضاً إذا كان الحصول على قيم دقيقة للمشاهدات غير ميسور وإن كان من الممكن ترتيبها.

(٣) إن مجموع رتب العينتين أى ر١ + ر٢ = مجموع الـ ن١+ ن٢ الأولى من سلسلة الأعداد الطبيعية وتساوى $\frac{(\text{ن١ ن٢})\ (\text{ن١+ن٢+١})}{\text{٢}}$ ومن هذه العلاقة فإنه يمكن معرفة ر١ من حساب ر٢ والعكس صحيح ولذلك يمكن إجراء الاختبار على أساس المقياس الإحصائى U لأى من العينتين خاصة إذا كانت ن١ = ن٢ وإلا فإنه يفضل - تسهيلاً للعمليات الحسابية - حساب U لأصغر العينتين حجماً.

(٤) إن مجموع $U_1 + U_2$ = ن١ن٢ وفى الحقيقة فإن توزيع العينات لـ U متماثل حول ن١ن٢ ÷ ٢ .

(٥) تستبدل الرتب المكررة بالوسط الحسابى لها ولا يؤثر ذلك على قيمة المقياس U^* ما لم تكن الرتب المكررة تشمل مجموعتى البيانات المقارنة. وفى الحالة الأخيرة فإنه قد ينصح باستخدام تصحيح معين يؤدى إلى زيادة محدودة فى قيمة ى* يمكن إهمالها.

(٥) اختبارات ولكوكسن التى تعتمد على الرتب:

١- اختبار ولوكسن لمجموع الرتب: اختبار W:

The Wilcoxon Rank-Sum Test

وهو اختبار مان-ويتنى لمجموع الرتب يستخدم مجموع الرتب لاختبار ما إذا كان مجتمعين مستقلين مستمرين غير مختلفين. أم أن أحدهما يختلف عن الآخر وذلك اعتماداً على بيانات عينتين مستقلتين تسحب كل من أحد المجتمعين والثانية من المجتمع الآخر وهو بالتالى يعتبر البديل اللابارامترى لاختبار ت للفرق بين متوسطين. وينسب هذا الاختبار إلى Frank Wilcoxon (١٨٩٢-١٩٦٥) وسوف نطلق عليه اختبار W كما يعرف به فى المراجع الإحصائية. وهناك جداول خاصة تستخدم فى حالة العينات الصغيرة ولكننا بدافع السهولة سوف نقتصر على تقديم هذا الاختبار فى حالة العينات الكبيرة ($ن_١$ ، $ن_٢$ كلاهما > ١٠). وخطوات الاختبار هى ذاتها لاختبار U خطوات الاختبار:

(١) تدمج مفردات العينتين المستقلتين (أ ، ب) مرتبة ترتيباً تصاعدياً مع تمييز رتب كل من العينتين مع إعطاء المفردات ذات نفس الترتيب الوسط الحسابى للرتب المكررة.

(٢) تجمع رتب كل من العينتين $د_أ$ ، $د_ب$.

(٣) واعتماداً على مجموع رتب أى من العينتين ، وإذا كان حجم كل من العينتين لا يقل عن ١٠ مفردات فإنه يمكن استخدام اختبار W الآتى:

W : مجموع رتب العينة أ أى $د_أ$.

وإذا صح فرض العدم بأن المجتمعين متماثلين وغير مختلفين فإن W أو $د_ا$ لها توزيع عينات توقعه:

$$\mu_W = \mu_{د_ا} = \frac{ن_ا (ن_ا + ن_ب + ١)}{٢} \qquad (١٩)$$

وتباينه:

$$\mu^٢_W = \sigma^٢_{د_ا} = \frac{ن_ا ن_ب (ن_ا + ن_ب + ١)}{١٢} \qquad (٢٠)$$

وبالتالي فإن أداة الاختبار هي:

$$ى = \frac{W - \mu_W}{\sigma_W} = \frac{د_ا - \mu_{د_ا}}{\sigma_{د_ا}} \qquad (٢١)$$

وهــذا المتغير العشوائي له توزيع ى : م(صفر ، ١) وبالتالي يمكن إجراء الاختبار في الاتجاه واحد أو اتجاهين وفق ما يمليه الفرض البديل وذلك بالمقارنة بقيم ى المعنوية كالمعتاد.

مثال (١٦):

استخدم بيانات مثال (٥-١٥) لاختبار الفرض فإن مجتمع المذكور ومجتمع الإناث كلاهما متماثلان عند α = ٥٪ .

الحـل :

$د_ا$ = W = ٣٣٣

$$\mu_{د_ا} = \mu_W = \frac{١٥ \ (١٥+٢٥+١)}{٢} = ٣٠٧,٥$$

$$\sigma^٢_{د_ا} = \sigma^٢_W = \frac{١٥\times٢٥ \ (١٥+٢٥+١)}{١٢} = ١٢٨١,٢٥$$

$$\sigma_{د_ا} = \sigma_W = ٣٥,٧٩٤٦$$

الفرض : متوسط القدرات الميكانيكية في المجتمعين متساويان

(أي أن $\mu_ا = \mu_ب$).

الفرض البديل : $\mu_{ا} \neq \mu_{ب}$ أى أنه إما $\mu_{ا} < \mu_{ب}$ أو $\mu_{ا} > \mu_{ب}$.

$ن_{ا}$ = ١٥، $ن_{ب}$ = ٢٥ وبالتالى يمكن استخدام اختبار العينات الكبيرة ى حيث:

$$ى = \frac{د_{ا} - \mu_{د_{ا}}}{\sigma_{د_{ا}}}$$

وإذا صح فرض العدم فإن المتغير العشوائى له توزيع ى . ويرفض فرض العدم إذا كانت $ى^*$ < -١,٩٦ أو > ١,٩٦ عند α = ٥٪ .

الحل:

ومن البيانات:

$$ى^* = \frac{٣٠٧,٥-٣٣٣}{٣٥,٧٩٤٦} = -٠,٧١٢ > -١,٨٦ \text{ أو } < ١,٩٦$$

∴ يقبل الفرض بتماثل المجتمعين من حيث القدرات الميكانيكية المتوسطة.

هذا وإذا أخذت $د_{ب}$ فى الاعتبار للاختبار بدلاً من $د_{ا}$ فإن:

$$د_{ب} = W = ٣٣٣$$

$$\mu_{د_{ب}} = \mu_{W} = \frac{ن_{ب}(ن_{ب} + ن_{ا} + ١)}{٢} = ٥١٢,٥$$

$$\sigma_{د_{ا}} = \sigma_{د_{ا}} = ٣٥,٧٩٤٦$$

$$∴ ى^* = \frac{٥١٢,٥-٤٨٧}{٣٥,٧٩٤٦} = -٠,٧١٢$$

ويقبل الفرض العدمى أيضاً ، هو ما كان متوقعاً إذ أن الاختبار فى إتجاهين وبالتالى فالفرض البديل واحد فى الحالتين.

كمـا يلاحـظ أنـنا وصـلنا إلى ذات القيمة لـ $ى^*$ وبالتالى نفس القرار باستخدام اختبار U:

ب: اختبار ولكوكسن لرتب الفروق بالإشارات: اختبار ر أو (T):

The Wilcoxon Signed-Rank Test

ويعمـــد هـــذا الاختبار على إشارات فروق الرتب وهو فى هذا يشبه اختبار الإشــــارة للقــــراءات المــــزدوجة ولكنه لاستخدامه لقدر أكبر من المعلومات عن اختــبار الإشــــارة حيــث يدخل فى الاعتبار ليس إتجاه الفروق فقط ولكن حجم الفـــروق بيـــن القــــراءات المزدوجة أيضاً ، لذلك فإنه أكفأ من اختبار الإشارة وبالتالى يفضل استخدامه بدلاً من اختبار الإشارة للفروق بين القراءات المزدوجة وسيتضح ذلك من المثال التالى:

خطوات الاختبار:

(١) تحسب الفروق بين كل زوج من أزواج القراءات المتناظرة $ف_ر$ (ر = ١ ، ٢ ،... ، ن) مـــع اســـتبعاد الفروق التى تساوى الصفر ويخفض حجم العينة مقابل كل زوج مستبعد من المشاهدات.

(٢) ترتـــب الفــــروق المطلقـــة $|ف_ر|$ تصاعدياً ثم يرصد لكل رتبة الإشارة الخاصة بالفرق الذى تناظره هذه الرتبة.

(٣) تجمـــع الرتـــب أخذاً فى الاعتبار الإشارات وإذا رمزنا إلى هذا المجموع بالرمز ر (وقد استخدم ولكوكسن الرمز T للدلالة على هذا المجموع).

(٤) وإذا كان حجم العينة ن ≤ ١٠ ومتى صح فرض العدم أى تساوى توزيع الظاهـــرة فـــى المجتمعين فإنه يتوقع أن يكون مجموع الفروق بإشارات موجـــبة = مجموع الفروق بإشارات سالبة ، وبالتالى فإن هذا المتغير له توزيع عينات توقعه:

$$\mu_ر = صفر \qquad (٢٢)$$

وتباينه:

$$\sigma^2_ر = \frac{ن (ن + ١) (٢ن + ١)}{٦} \qquad (٢٣)$$

وبالتالى فإن المتغير العشوائى:

$$ى = \frac{ر - صفر}{\sqrt{\frac{ن\,(ن+1)\,(2ن+1)}{6}}} \qquad (24)$$

له توزيع ى : م (صفر ، ١)

واعتماداً على ذلك يمكن إجراء الاختبار بتساوى المجتمعين المتناظرين.

مثال (١٧):

بعد أن أدخلت وزارة البترول تعديلاً على البنزين وأصبح البنزين فى بعض المحافظـــات خالياً من الرصاص ولاختبار ما إذا كان إستهلاك البنزين / اللتر لم يخـــتلف بعد إدخال هذا التعديل عنه قبل ذلك فقد أجريت تجربة النوعين عل ى عينة من ١٥ سيارة مختلفة الطراز وعلى نفس الطريق باستخدام البنزين الخالى من الرصاص أ ومرة أخرى بعد استخدام البنزين العادى ب ، وقد تم تخصيص نو ع البنزين عشوائياً على السائقين ، ومن خصص له النوع أ فى الدورة الأولى خصص له النوع ب فى الدورة الثانية وبالعكس. وكانت نتيجة التجربة كالآتى:

طراز السيارة	كيلومتر / اللتر		إشارة ف = أ−ب	ف	|ف|	رتبة $\|ف_ر\|$	الترتيب بالإشارات
	بنزين خالى من الرصاص (أ)	بنزين عادى (ب)					
١	٢٢,١	٢٣,٧	−	١,٦−	١,٦	٩	٩−
٢	١٥,٧	١٦,١	−	٠,٤−	٠,٤	٣	٣−
٣	١٨,٢	١٩,٠	−	٠,٨−	٠,٨	٥	٥−
٤	١٩,٠	١٩,٠	صفر	صفر	صفر	ــــ	ــــ
٥	٢٥,٧	٢٤,٢	+	١,٥+	١,٥	٧,٥	٧,٥+
٦	١٩,٢	٢٢,٠	−	٢,٨−	٢,٨	١٢,٥	١٢,٥−
٧	١١,٩	١٢,١	−	٠,٢−	٠,٢	١	١−

٨	٢٨,٠	٣٠,٢	-	٢,٢-	٢,٢	١٠,٥	١٠,٥-
٩	٣٥,٠	٣٧,٨	-	٢,٨-	٢,٨	١٢,٥	١٢,٥-
١٠	٢٧,٠	٣٠,٠	-	٣,٠-	٣,٠	١٤	١٤-
١١	١٩,٧	١٩,٠	+	٠,٧+	٠,٧	٤	٤+
١٢	١٦,٣	١٥,١	+	١,٢+	١,٢	٦	٦+
١٣	٢١,٢	٢٣,٤	-	٢,٢-	٢,٢	١٠,٥	١٠,٥-
١٤	٢٢,٢	٢٠,٧	+	١,٥+	١,٥	٧,٥	٧,٥+
١٥	٢٧,٨	٢٨,١	-	٠,٣-	٣,٠	٢	٢-

مجـ ر = -٥٥

الفرض العدمى : متوسط المسافة المطوعة (كيلومتر / باللتر) باستخدام نوعين البنزين أ ، ب واحد.

الفرض البديل : إن متوسط المسافة المقطوعة (كيلومتر / لتر) فى حالة استخدام البنزين النوع أ أعلى منه باستخدام النوع ب .

مستوى المعنوية α = ٥٪ .

باستخدام اختبار الإشارة س = ٤ .

ح (س : + ≥ ٤ | ن = ١٤)

= ٠,٠٠١ + ٠,٠٠٦ + ٠,٠٢٩ + ٠,٠٩ = ٠,١٢٦ > ٠,٠٥

لذلك يقبل فرض العدم باستخدام اختبار الإشارة.

وباستخدام اختبار ولكوكسن للرتب بالإشارات:

$$\sigma_{ر} = \sqrt{\frac{١٤(١٤+١)(٢٨+١)}{٦}} = ٣١,٨٦$$

$$\therefore ى^{*} = \frac{-٥٥}{٣١,٨٦} = -١,٧٢٦ < ١,٦٤٥$$

لذلك يقبل فرض العدم باستخدام اختبار ولكوكسن للرتب بالإشارات ويمكن أن يجرى الاختبارين كل فى اتجاهين أى أنه باستخدام اختبار الإشارة

ح (س : + ≥ ١.٤ | ن = ٤) = ٠,١٢٦ × ٢ = ٠,٢٥٢ > ٠,٠٢٥

وبالـــتالى يقــبل فــرض العــدم. وكذلك بالنسبة لاختبار ولكوكسن للرتب بالإشارات:

ى* = -١,٧٢٦ > -١,٩٦ > ١,٩٦ أى يقبل فرض العدم أيضاً.

وهناك مدخل آخر لاختبار ولكوكسن للرتب بالإشارات - كما أوضحنا فى الفقرة السابقة ويعتمد فى ذلك على توزيع مجموع الرتب الموجبة $ر_+$ أو (T_+) أو مجموع الرتب السالبة $ر_-$ أو (T_-) ، وإذا صح فرض العدم فإنه يتوقع أن تكون $ر_+ = ر_-$. أى أن فرق بين أزواج القيم المتناظرة له نفس الاحتمال لأن يكـــون فـــرقاً موجباً أو سالباً ، وبالتالى فإن المتغير العشوائى $ر_+$ أو $(ر_-)$ له توزيع عينات توقعه:

$$\mu_{ر_+} \text{ أو } (\mu_{ر_-}) = \frac{ن\,(ن+١)}{٤} \qquad (٢٥)$$

وتباينه:

$$\sigma^٢_{ر_+} \text{ أو } (\sigma^٢_{ر_-}) = \frac{ن\,(ن+١)\,(٢ن+١)}{٢٤} \qquad (٢٦)$$

وبالتالى فإن المتغير العشوائى:

$$ى = \begin{cases} \dfrac{ر_+ - \mu_{ر_+}}{\sigma_{ر_+}} \\ \text{أو} \\ \dfrac{ر_- - \mu_{ر_-}}{\sigma_{ر_-}} \end{cases} \qquad (٢٧)$$

له توزيع ى : م (صفر ، ١) واعتماداً على ذلك يمكن إجراء الاختبار على أساس $ر_+$ أو $(ر_-)$.

مثال (١٨):

وتطبيقاً لذلك باستخدام $ر_+$ وباستخدام بيانات المثال السابق (٥-١٧) فإن:

$ر_+$ = ٢٥ $\mu_{ر_+}$ = ٥٢,٥

$\sigma_{ر_+}$ = ١٥,٩٣

ويستكمل اختبار الفرض كالآتي:

الفرض العدمي : $\mu_أ = \mu_ب$

الفرض البديل : $\mu_أ > \mu_ب$

α = ٠,٠٥

ومن البيانات:

$$ى^* = \frac{٢٥-٥٢,٥}{١٥,٩٣} = -١,٧٢٦ < ١,٦٤٥$$

ويقبل فرض العدم: إذا كان الاختبار في اتجاه واحد عند α = ٥٪ كما يقبل الفرض العدمي إذا كان الاختبار في اتجاهين حيث -١,٩٦ > ١,٧٢٦ < ١,٩٦ عند α = ٥٪ .

وإذا طبق ذات الاختبار على مجموع الرتب السالب $ر_-$ أو (T_-) .

$ر_-$ أو (T_-) = ٨٠

$\mu_{ر_-}$ أو (μ_{T_-}) = ٥٢,٥

$\sigma_{ر_-}$ أو (σ_{T_-}) = ١٥,٩٣

$$\therefore ى^* = \frac{٢٧,٥}{١٢,٩٣} = ١,٧٢٦$$

وإذا كان الفرض العدمي $\mu_أ = \mu_ب$

مقابل الفرض البديل $\mu_أ \neq \mu_ب$ في اتجاهين.

ى* = ١,٧٢٦ > -١,٩٦ ، < ١,٩٦ لذلك يقبل الفرض العدمى ، أما إذا كـان الفرض البديل هو أن التوزيع الاحتمالى لـ أ قد إنحاز إلى الجانب الأيسر من التوزيع الاحتمالى لـ ب أى أن $\mu_أ < \mu_ب$ وهو اختبار طرف أيسر ولأن ى* = ١,٧٢٦ > -١,٦٤٥ لذلك يقبل فرض العدم.

ملحوظة:

قـد لا يتفق القرار بالنسبة للفرض العدمى فى حالة استخدام اختبار الإشارة مـع القـرار باستخدام اختبار الرتب بالإشارات لولكوكسن لأن الاختبار الأخير يسـتخدم معلومـات أكثر من البيانات وفى حالة كهذه يفضل استخدام الاختبار الأخير الأكثر تمييزاً.

(٦) تحليل التباين فى اتجاه واحد بالرتب / اختبار كروسكال -والس
Kruskal-Wallis One-way Analysis of Variance

سـبق أن قدمـنا اختبار تجانس توزيع ظاهرة فى عدة مجتمعات باستخدام توزيـع كـا٢ كأسلوب الاختبار فى حالة تعدد العينات المستقلة. وفى هذه الفقرة سوف نقدم أسلوباً آخر هو تحليل التباين فى اتجاه واحد باستخدام الرتب.

اختبار كروسكال/والس:

وهـو كالتحليل فى إتجاه واحد باستخدام تحليل التباين - أى التصميم كامل العشـوائية - يهدف إلى اتخاذ قرار فى شأن عدة عينات مستقلة تسحب من عدة مجتمعات وأنها لا تختلف فى توزيع المتغير داخلها أى ما إذا كانت الفروق بين متوسطات العينات تعكس فروقاً بين المجتمعات التى سحبت منها تلك العينات أم لا ؟

ويعـتمد هـذا الاختبار على الرتب خاصة إذا كانت القياسات ترتيبية أو أن الـتوزيع المعتاد غير محقق وإذا صح فرض العدم بأن متوسطات (أو الوسيط) للمجـتمعات التى سحبت منها العينات متساوية فإن متوسطات الرتب لكل عينة

ستكون متساوية وإذا جمعت الرتب فى جميع العينات أو فى العينة التجميعية فإن متوسطات الرتب = $\frac{ن+1}{2}$ وبالتالى فإن القيمة المتوقعة للفروق بين متوسط الرتب لكل عينة والمتوسط العام للرتب التجميعية أى توقع

$$\left(\frac{ر_و}{ن_و} - \frac{ن+1}{2}\right) = \text{صفر}$$

وجدير بالذكر أن شكل التوزيـــع فى المجتمعـــات غير مطلوب التحقق منه وكل ما يتعيـــن

توافره هو استقلال العينات المسحوبة من المجتمعات المختلفة.

خطوات الاختبار:

(١) يحـــدد الفـــرض العدمـــى وهـــو أنه لا يوجد فرق بين توزيع المتغير فى المجتمعات موضوع الاختبار أما الفرض البديل فهو أن التوزيعات متباينة.

(٢) ويحدد مستوى المعنوية α .

(٣) ترتب القياسات فى جميع الـــ و = ١ ، ٢ ، ... ، م عينة – والتى سحبت كل من مجتمع الدراسة – مجتمعة وترصد الرتب المناظرة لتلك القياسات مع مراعاة الترتيب التصاعدى ثم تجمع الرتب $ر_و$ فى كل عينة.

وتستخدم أداة الاختبار:

$$هـ = \frac{12}{ن(ن+1)} \text{مجـ}\, ن_و \left\{\left(\frac{ر_و}{ن_و} - \frac{ن+1}{2}\right)^2\right\} \qquad (28)$$

وهـــى صـــيغة تماثل كا^2 لجودة المطابقة والتى ذكرت فى بداية هذا الفصل ويقترب توزيعها من توزيع كا^2 أو الصيغة (٢٩) التالية:

$$هـ = \frac{12}{ن(ن+1)} \sum_{و=1}^{م} \frac{ر_و^2}{ن_و} - 3(ن+1) \qquad (29)$$

حيث و = ١ ، ٢ ، ... ، م عينة $ن_و$: حجم العينة و

ن = مجـ ن و   ر و = مجموع الرتب فى العينة و

حيـــث يتبع المتغير العشوائى هـ توزيع كا٢ بدرجات حرية م − ١ متى كان حجم كل العينة ن و > ٥ . أما إذا كان عدد العينات م = ٣ ، وحجم العينة ن و ≤ ٥ فـــى كل عينة فإن تقارب توزيع هـ لتوزيع كا٢ يصبح غير محقق ويرجع إلى ملحق (١١) الذى يستخدم المعادلة السابقة (٢٨) لحساب احتمال مشاهدة قيمة لـــ هـ ≤ القـيمة المحسوبة هـ* لقيم مختلفة لـــ ن و . وعليه فيرفض الفرض العدمى إذا كانت ح (هـ ≤ α) ويقبل فيما عدا ذلك.

وكمـــا تبيـــن من الفقرة السابقة فإن هذا الاختبار يحدد ما إذا كانت مجاميع الرتـــب متباينة وبالتالى فإن العينات لا يتوقع أن تكون من مجتمع واحد (أو عدة مجتمعات متجانسة).

مثال (١٩):

اســـتخدم أسلوب التحليل فى اتجاه واحد بالرتب (اختبار كروسكال/والس) لتحليل بيانات مثال (٣−٢) بالفصل الثالث.

الحـل:

الآلة								مجـ ر و
	(١)	س١	٤٧	٥٣	٤٩	٥٠	٤٦	
		الرتب ر١	٢	١٠	٣,٥	٥,٥	١	٢٢
الآلة	(٢)	س٢	٥٥	٥٤	٥٨	٦١	٥٢	
		الرتب ر٢	١٣	١١,٥	١٤	١٥	٩	٦٢,٥
	(٣)	س٣	٥٤	٥٠	٥١	٥١	٤٩	
		الرتب ر٣	١١,٥	٥,٥	٧,٥	٧,٥	٣,٥	٣٥,٥

الفـــرض العدمـــى : م١ = م٢ = م٣ أى أن الآلات الثلاث لا تختلف فى إنتاجيتها.

α = ٠,٠٥ وباستخدام وسيلة الاختبار:

$$\text{هـ} = \frac{١٢}{\text{ن}(\text{ن}+١)} \sum_{\text{و}=١}^{\text{م}} \frac{\text{ر}_{\text{و}}^{٢}}{\text{ن}_{\text{و}}} - ٣(\text{ن}+١)$$

$$\text{هـ}^{*} = \frac{١٢}{١٥ \times ١٦} \left\{ \frac{٢٢^{٢}}{٥} + \frac{٦٢,٥^{٢}}{٥} + \frac{٣٥,٥^{٢}}{٥} \right\} - ٣ \times ١٦$$

$$= \frac{٠,٠٥}{٥} (٤٨٤ + ٣٩٠٦,٢٥ + ١٢٦٠,٢٥) - ٤٨$$

$$= ٥٦,٥٠٥ - ٤٨ = ٨,٥٠٥$$

وباستخدام المعادلة (٢٨) فإن:

$$\text{هـ}^{*} = \frac{١٢}{١٥ \times ١٦} \times \left\{ \left(\frac{٢٢}{٥} - ٨\right)^{٢} + \left(\frac{٦٢,٥}{٥} - ٨\right)^{٢} + \left(\frac{٣٥,٥}{٥} \; ٨\right)^{٢} \right\}$$

$$= (٣٤,٠٢) = ٨,٥٠٥$$

وهى نفس القيمة السابقة.

ومن ملحق (١١)

ح ($\text{هـ} \geq ٨$) أقل من ٠,٠٠٩ لذلك يرفض الفرض اعدمى عند $\alpha = ٠,٠٥$

هذا ويلاحظ أننا لو طبقنا تقارب توزيع هـ لتوزيع كا^{٢} فإن:

$\text{كا}^{2*} = \text{هـ}^{*} = ٨,٥٠٥ > \text{كا}^{٢}_{٠,٠٥ ، ٢} = ٥,٩٩$ فيرفض الفرض أيضاً.

كمـــا سبق أن رفضنا الفرض باستخدام أسلوب تحليل التباين (راجع الفصل الأول).

ملحوظة:

إذا كانت بعض الرتب مكررة - كما هو الحال فى المثال الأخير - فتعطى القيم المكررة الوسط الحسابى للرتب المكررة ثم تصحح هـ* كالآتى:

$$\text{هـ}^{*} \text{ مصححة} = \frac{\text{هـ}}{١ - \frac{\text{مجـ ث}}{\text{ن}^{٣} - \text{ن}}} \qquad (٣٠)$$

حيث $\text{ث} = \text{ت}^{٢} - \text{ت}$ ، ت = عدد المشاهدات المكررة فى كل عينة.

وهذا التصحيح يؤدى إلى زيادة قيمة هـ* وبالتالى فإنه يفقد قيمته العملية إذا كانت قيمة هـ* غير المصححة تؤدى إلى رفض الفرض العدمى.

وبالإضافة إلى ذلك فإن كثرة عدد الرتب المكررة يقلل من دقة هذا الاختبار الذى يقوم على فرض إستمرارية التوزيعات.

ملاحظات ختامية على الاختبارات التى تعتمد على عدة عينات مستقلة:

سبق أن قدمنا فى الفصل الثالث تحليل التباين كمدخل للتحليل فى حالة تعدد العينات. كما قدمنا فى هذا الفصل اختبار كا² لتجانس توزيع ظاهرة فى عدة مجتمعات كمدخل آخر للاختبار فى حالة تعدد العينات وكان المدخل الأول بارامترى والثانى مدخل لابارامترى ثم عرضنا لمدخل لابارامترى ثان هو تحليل التباين فى اتجاه واحد بالرتب لكروسكال والس للاختبارات فى العينات المتعددة ويعنينا أن تعرف على السمات الرئيسية التى تميز كل من هذه المداخل الثلاث:

(١) يلاحظ أننا استخدمنا للاختبارات اللابارامترى لم نفترض الشروط الخاصة بتحليل التباين – راجع الفصل الثالث – من حيث التوزيع الإعتدالى للمتغيرات فى المجتمعات واستقلال الأخطاء وتوزيعها الإعتدالى و ... إلخ إكتفاء بفرض توافر شروط أيسر كاستمرارية التوزيع أو إستقلالية العينات المسحوبة من مجتمعات مختلفة.

(٢) يستخدم تحليل التباين فى اتجاه واحد لكروسكال – والس متى كان توزيع المتغير مستمراً والقياسات من النوع الترتيبى على الأقل وكفاءته النسبية بالمقارنة بتحليل التباين ٩٥,٥٪ مما يجعل استخدام اختبار كروسكال/والس مفضلاً عندما نتشكك فى عدم تحقق فى بعض الفروض الخاصة بتحليل التباين البارامترى مثل عدم تماثل توزيع المتغير فى المجتمعات أى تساوى تباينها.

(٣) أما إذا كانت القياسات من النوع الوصفى فلا بديل عن اختبار كا².

(٧) معامل إرتباط الرتب لسبيرمان واختبار معنويته:

سبق أن عرضنا معامل إرتباط الرتب لسبيرمان والمعرف كالآتي:

$$ر = ١ - \frac{٦ مجـ ف^٢}{ن (ن^٢ - ١)} \qquad (٣١)$$

والـــذى يســـتخدم لقياس العلاقة الخطية بين متغيرين س ، ص حيث ف = رتبة س-رتبة ص لجميع أزواج القيم المتناظرة للمتغيرين س ، ص.

وكمـــا أوضـــحنا فإنـــه يمكـــن أن يستخدم هذا المعامل لقياس العلاقة بين المتغيريـــن إذا كانت القياسات ترتيبية وفى هذه الحالة لا يصلح معامل الإرتباط الخطـــى البســـيط لبيرســـون لقياس تلك العلاقة ، كما يمكن استخدامه إذا كانت القياسات يمكن أن تستخدم فى ترتيب قيم س ، ص تصاعدياً أو تنازلياً.

كمـــا أوضحنا فى الفصل الرابع من هذا الكتاب كيفية اختبار معنوية معامل الارتـــباط الخطى البسيط لبيرسون إعتماداً على اختبار معنوية معامل الإنحدار الخطى البسيط β_0 = صفر حيث اختبار β_0 يعتمد على أداة الاختبار:

$$\frac{\text{متوسط مربعات للإنحدار الخطى البسيط}}{\text{متوسط مربعات البواقى}} = \frac{ر^٢ (ن - ٢)}{١ - ر^٢} \qquad (٣٢)$$

وهـــذا المتغير العشوائى له توزيع **ف** بدرجات حرية (١ ، ن - ٢) وجذره التربيعى:

$$\frac{ر \sqrt{ن - ٢}}{\sqrt{١ - ر^٢}} \qquad (٣٣)$$

لـــه توزيـــع **ت** بدرجـــات ن - ٢ وهـــذا الأخير هو أداة الاختبار للفرض الإحصائى المتعلق بمعامل الإرتباط الخطى البسيط لبيرسون ρ_0 = صفر مقابل الفرض البديل $\rho \neq$ صفر وعليه فإن اختبار β_0 هو نفسه اختبار ρ_0 .

وفـــى هــذه الفقـــرة سوف نقدم كيفية اختبار معنوية معامل ارتباط الرتب لسبيرمان.

أ: معامل إرتباط الرتب لسبيرمان

مثال (٣٠):

لدراســـة العلاقـــة بين طول مدة التغيب وعدد الوحدات المعيبة فقد أجريت داســـة جمعت من خلالها بيانات عن معدل التغيب الأسبوعى المتوسط (س) فى عينة مكونة من ١٢ أسبوعاً وعدد الوحدات المعيبة المنتجة (ص) خلال الأسابيع الاثنتى عشرة فكانت النتائج كما هو فى جدول (١٣) التالى:

جدول (١٣)

معدل التغيب المتوسط وعدد الوحدات المعيبة فى عينتين

الأسبوع	١	٢	٣	٤	٥	٦
معدل التغيب	٧,٣	٦,٤	٦,٢	٥,٥	٦,٥	٤,٧
عدد الوحدات المعيبة المنتجة	٢٢	١٧	٩	٨	١٢	٥
الأسبوع	٧	٨	٩	١٠	١١	١٢
معدل التغيب	٥,٨	٧,٩	٦,٧	٩,٦	١٠,٣	٧,٢
عدد الوحدات المعيبة المنتجة	٧	١٩	١٣	٢٩	٣٣	١٨

أوجد معامل ارتباط الرتب بين س: معدل التغيب الأسبوعى المتوسط ، ص: عدد الوحدات المعيبة المنتجة واختبر معنوية معامل الارتباط.

الحـل:

الأسبوع	س معدل التغيب	ص عدد الوحدات	$ر_س$	$ر_ص$	ف	$ف^٢$
١	٧,٣	٢٢	٩	١٠	١-	١
٢	٦,٤	١٧	٥	٧	٢-	٤
٣	٦,٢	٩	٤	٤	صفر	صفر
٤	٥,٥	٨	٢	٣	١-	١
٥	٦,٥	١٢	٦	٥	١	١
٦	٤,٧	٥	١	١	صفر	صفر
٧	٥,٨	٧	٣	٢	١	١
٨	٧,٩	١٩	١٠	٩	١	١
٩	٦,٧	١٣	٧	٦	١	١
١٠	٩,٦	٢٩	١١	١١	صفر	صفر
١١	١٠,٣	٣٣	١٢	١٢	صفر	صفر
١٢	٧,٢	١٨	٨	٨	صفر	صفر
					صفر	١٠

$$ر = ١ - \frac{٦ مجـ ف^٢}{ن(ن^٢-١)} = ١ - \frac{٦ \times ١٠}{١٢ \times ١٤٣}$$

$$= ١ - \frac{٦٠}{١٧١٦} = ١ - ٠,٠٣٥ = ٠,٩٦٥$$

ب: اختبارمعنوية معامل ارتباط الرتب لسبيرمان:

ولإختبار معنوية معامل الإرتباط ρ = صفر

الفرض العدمى ρ = صفر والبديل $\rho \neq$ صفر وهو اختبار فى اتجاهين.

وبالــرجوع إلى ملحق (١٢) والذى يحدد الحدين الأدنى والأعلى (تقدير ρ بفــترة ثقة) لــ ρ لعدد من أزواج القيم ن $\leq$ ٣٠ لمستويات معنوية مختلفة نجد أنه عند α = ١٪ فإن:

-٠,٨١٨٢ $\leq$ ρ $\leq$ ٠,٨١٨٢ ولذلــك يــرفض الفــرض العدمى إذ أن ر = ٠,٩٦٥ تقع خارج منطقة قبول الفرض بعدم وجود علاقة بين س ، ص.

وفى حالة العينات الكبيرة أى لعدد من أزواج القيم > ٣٠ فإنه يمكن الإعتماد على توزيع العينات لمعامل إرتباط الرتب وهو توزيع يؤول إلى التوزيع المعتاد توقعه (صفر) وتباينه $\frac{1}{ن - 1}$ لذلك فإنه يمكن استخدام المتغير العشوائى:

$$ى = \frac{ر - صفر}{\frac{1}{\sqrt{ن - 1}}} \sim م\ (صفر ، ١) \qquad (٣٤)$$

لاختبار الفروض الإحصائية المتعلقة بـ ρ .

مثال (٣١):

فى دراسة اجتماعية عن العلاقة بين مستوى الذكاء للزوج والزوجة فى عينة عشــوائية مكونــة مــن ٣٢ مــن حالات الزواج وجد أن معامل إرتباط الرتب لسبيرمان كان = ٠,٨٠٩ اختبر معنوية هذه العلاقة عند α = ٥٪ .

الحل:

الفرض العدمى $\rho_.$ = صفر

الفرض البديل ρ $\neq$ صفر

α = ٥٪

ولأن ن = ٣٢ لذلك يمكن استخدام التقارب الإعتدالى لإختبار الفرض حيث:

$$\text{ى} = \frac{\text{ر} - \rho_0}{\frac{1}{\sqrt{\text{ن} - 1}}}$$

ويرفض الفرض العدمى إذا كانت $\text{ى}^{*} < -١,٩٦$ أو $> ١,٩٦$

ومن بيانات العينة $\text{ى}^{*} = \frac{٠,٨٠٩ - \text{صفر}}{\frac{١}{\sqrt{٣١}}} = ٠,٨٠٩ \times ٥,٥٦٨ = ٤,٥٠٤$

لذلــك يــرفض الفرض العدمى والعلاقة بين مستوى ذكاء الزوج والزوجة علاقة معنوية.

تمارين

١- ألقيت ٣ قطع عملة ٢٠٠ مرة ورصد المتغير العشوائى س الدال على عدد مرات ظهور الكتابة فى كل رمية مستقلة (س = ٠ ، ١ ، ٢ ، ٣) فكانت النتائج كالآتى:

س	٠	١	٢	٣
ك	٣٤	٦١	٧٢	٧٣

فهل تدل نتائج هذه التجربة على أن قطع العملة غير متحيزة عند مستوى المعنوية ٥٪ ؟

٢- الجدول التالى يوضح توزيع أفراد عينة عشوائية من المتقدمين لشغل بعض الوظائف القيادية بحسب الوقت الذى استغرق فى أداء اختبار للصلاحية:

الزمن بالدقيقة	عدد الأفراد
٢٤ دقيقة أو أقل	١٥
٢٥-	٥٠
٣٠-	٧٥
٣٥-	٤٠
٤٠-	١٥
٤٥ دقيقة فأكثر	٥

اختبر عند α = ٥٪ أن هذه العينة سحبت من مجتمع معتاد μ = ٣٢,١ دقيقة ، σ = ٥,٦

٣- تيسيراً على عملاء أحد المصارف لصرف شيكات من حساباتهم عن طريق شبابيك خاصة لراكبى السيارات drive-in وحتى يمكن للمصرف أن يحدد عدد الشبابيك التى تتفق مع الطلب على صرف شيكات عن هذا

الطريق لتحقيق السيولة المناسبة فقد رصد المصرف فى أحد فروعه عدد مرات تقدم العمــلاء بشيكــات للصرف فى الدقيقة وكانت نتيجة هذه التجربة فى ٧٥٠ حالة كالآتى:

عــدد مــرات وصــول العميل بسيارته لصرف شيك فى الدقيقة	صفر	١	٢	٣	٤	٥	٦	٧+
عدد العملاء	٤٦	٢٥٠	٢٦٠	١٠٠	٥٠	٣٤	١٠	صفر

والمــراد اختبار أن هذا التوزيع يتبع التوزيع الاحتمالى المعروف باسم " توزيع بواسون " عند مستوى المعنوية ٥٪ .

٤- لدراســة العلاقة بين درجة نجاح العاملين فى عملهم بأحد المؤسسات وما حققــه كــل منهم فى برنامج تدريبى فقد سجلت البيانات التالية من عينة عشوائية من ٤٠٠ عامل.

درجة النجاح فى العمل	درجة النجاح فى البرنامج التدريبى		
	دون المتوسط	متوسط	أعلى من المتوسط
ضعيف	٢٣	٦٠	٢٩
متوسط	٢٨	٧٩	٦٠
جيد	٩	٤٩	٦٣

حدد الفرض الذى تختبره ثم أجرى الاختبار المناسب عند α = ٥٪

٥- لدراسة العلاقة بين طول مدة خدمة العامل فى أحد المتاجر الكبرى وتقبل هــؤلاء العاملين لأثر التخفيضات السنوية على حجم المبيعات فقد جمعت البــيانات التالــية مــن أفراد عينة عشوائية مكونة من ٤٤٠ عاملاً. حدد الاختبار الإحصائى المناسب ثم استخدمه عند مستوى المعنوية α = ٥٪ .

التقييم	مدة الخدمــــــــــــة			
	سنتين أو أقل	٣ — ٥	٦ — ١٠	أكثر من ١٠
تزاد المبيعات	٤٨	٥١	٤٣	٣٥
لا تتأثر	٢٢	٤٢	٥٩	٥٧
تنخفض	١٥	٢٢	٣٣	١٣

٦- الجــدول الــتالى يبين توزيع عينة من الأسر بحسب الدخل السنوى وعدد الأطفال بكل:

الدخل السنوى للأسرة بالجنيه	عدد الأطفــــــال			
	صفر	١	٢	أكثر من ٢
أقل من ٥٠٠	١٥	٢٧	٥٠	٤٣
٥٠٠ — ١٠٠٠	٢٥	٣٧	١٢	٨
أكبر من ١٠٠٠	٨	١٣	٩	١٠

بيــن الفــرض الإحصائى المناسب ثم استخدمه فى اختبار الدخل السنوى وعلاقته بعدد الأطفال عند α = ٥٪ .

ثم بفرض أن عينة الأسر ذات الدخل أقل من ٥٠٠ جنيه بلغت ١٣٥ أسرة وعينة الأسر ذات الدخل من ٥٠٠ إلى ١٠٠٠ جنيه كانت ١٨٢ أسرة أما عينة الأسر ذات الدخل أكبر من ١٠٠٠ جنيه فكانت مكونة من ٤٠ أسرة ماذا يكون الاختبار ؟

وما هو قرارك عند α = ٥٪ ؟

٧- الجــدول الــتالى يبين عدد الحوادث التى وقعت للعمال والتى سجلت فى فترات العمل الثلاث فى أحد الوحدات الإنتاجية:

فترة العمــــــل			
الأولى	الثانية	الثالثة	مجموع
٢٥	٣٠	٤١	٩٦

هل هناك شك فى عدم وجود اختلاف معنوى بين فترات العمل الثلاث من حيث عدد الحوادث التى تقع عند مستوى المعنوية α = ٥٪ .

٨- إذا كــان متوســط الوقت الذى تم فيه تجميع الوحدة من منتج ما فى أحد مراكــز التجمــيع هــو ٣٢ دقيقة بإنحراف معيارى قدره ٠,١٢٢ دقيقة. وتقضى خطة ضبط الإنتاج فى مركز التجميع أن يكون تباين وقت التجميع داخــل الفــترة (٠,٠٠٨ ، ٠,٠٣) دقيقة وإلا أوقفت عملية التجميع لإعادة المعايــرة. وقــد تبين من تسجيل وقت التجميع لعينة عشوائية عددها ٢٥ وحدة أن ع = ٠,١٧٠ دقيقة. عين التقدير بفترة ثقة ٩٥٪ لـ σ^2.

٩- اختــبر الفــرض بــأن σ^2 = ٠,٠١٥ دقيقة مقابل الفرض البديل σ^2 > ٠,٠١٥ اعتماداً على نتائج التجربة السابق بمستوى معنوية ٠,٠١ .

١٠- الجــدول التالى يبين نسبة الوحدات المعيبة من إنتاج آلة ما فى ٢٥ يوماً متتالية:

٨,٢ ١٠,٣ ٨,٩ ٩,٤ ١٠,٣ ٩,٤ ١٢,٣ ١٠,٠

٨,٤ ١١,٤ ١١,١ ١٢,٠ ١١,٨ ١٠,١ ٨،٨ ١٠,٤

٩,٣ ٩،٩ ١٢,٢ ٧,٤ ٨,٦ ٩,٧ ١٠,٩ ١١,٩

١١,٢١

اختــبر عــند مســتوى المعنوية ٥٪ أن نسبة المعيب > الوسيط = نسبة المعيب < الوسيط.

١١- البيانات التالية هى طول مدة الانتظار فى محطة الأتوبيس بالدقيقة لأفراد عينة عشوائية مكونة من ١٢ راكباً:

٢ ، ٧ ، ٩ ، ٣ ، ٧ ، ٩ ، ٥ ، ٧ ، ٨ ، ٧ ، ١٠ ، ٤

استخدم اختبار الإشارة عند α = ٥٪ لإختبار الفرض μ = ٥ دقائق مقابل الفرض البديل $\mu \neq$ ٥ دقائق ثم أعد الاختبار باستخدام اختبار بارامترى مناسب وعلق على نتائجك.

١٢- فيما يلى المدة (بالدقيقة) التى استغرقها أفراد عينتين عشوائيتين من الرجال والنساء فى أداء اختبار للمقابلة للإلتحاق بوظائف أحد المصارف الكبرى:

رجال ٦,٥ ١٠،٠ ٧,٠٠ ٩,٨ ٨,٥ ٨,٢ ٩,٠ ٨,٢
١٠,٩ ٨,٧ ٦,٧ ٨,١ ٧,٩ ٦,٤ ٨,٩

نساء ٨,٦ ٧,٨ ٨,٣ ٦،٦ ١٠,٥ ٦,٣ ٩,٣ ٨,٤
٩,٧ ٨،٨ ٩،٩ ٧,٦

إستخدم اختبار مجموع إشارات الرتب U لاختبار أن العينتين من مجتمعين متماثلين عند α = ٥٪ .

١٣- سجل عدد حوادث المرور الأسبوعية التى وقعت عند تقاطع ما فى إحدى المدن خلال السنة الماضية فكانت على النحو التالى:

عدد الحوادث	صفر	١	٢	٣	٤	٥	٦
التكرار	١٠	١٢	١٢	٩	٥	٣	١

(أ) ما هو التوزيع الاحتمالى الذى يصف هذه النتائج ؟ ولماذا كان اختيارك لهذا التوزيع؟

(ب) اختبر الفرض بأن النتائج المسجلة عن حوادث المرور عن تلك السنة. تتبع فى توزيعها التوزيع الاحتمالى الذى اخترته فى (أ) من هذا السؤال وذلك عند مستوى المعنوية α = ٥٪ .

١٤- فيما يلى عدد الوحدات المعيبة التى أنتجتها آلة ما فى ٢٤ يوم عمل متتالية:

١٥ ١١ ١٧ ١٤ ١٦ ١٢ ١٩ ١٧
٢١ ١٥ ١٧ ١٩ ٢١ ١٤ ٢٢ ١٦
١٩ ١٢ ١٦ ١٤ ١٨ ١٧ ٢٤ ١٣

اختبر عشوائية العينة عند α = ٥٪ .

١٥- الآتى بيان نتيجة فحص عينات زجاجية متابعة للتلف نتيجة الشحن والنقل (سليم / تالف : س / ت).

س / س / س / س / ت / س / ت / ت / ت / س

س / ت / ت / س / س / س / س / س / س / س

ت / ت / ت / ت / س /

اختبر عشوائية نتيجة فحص هذه الرسائل عند α = ١٪ .

١٦- الآتى بيان قيمة المنفق بالمليون جنيه على بحوث التطوير والتنمية فى ١٥ مؤسسة عامى ١٩٦٢ ، ١٩٧٢ .

اختبر الفرض بأنه لم يحدث تغير فى الإنفاق على بحوث التطوير والتنمية ما بين ١٩٦٢ ، ١٩٧٢ مقابل الفرض البديل أن الإنفاق سنة ١٩٧٢ قد زاد عما كان عليه سنة ١٩٦٢ عند مستوى المعنوية ٥٪ .

المؤسسة	١	٢	٣	٤	٥	٦	٧	٨
المنفق سنة ١٩٦٢	١٢	٨	١٤	٩	١٧	٢٠	٧	١١
١٩٦٢	١٤	٩	١٧	٨	٢٠	٢٢	١١	١١
تابع المؤسسة	٩	١٠	١١	١٢	١٣	١٤	١٥	
المنفق سنة ١٩٦٢	١٢	١١	٨	٥	١٣	٢٢	١٧	
١٩٧٢	١٥	١٣	١٠	١٠	١٧	٢٧	١٦	

١٧- فيما يلى سعر البيع بالقطاعى بالجنيه للوحدة من نوعين مختلفين من سلعة ما فى عينة من منافذ البيع فى بلد ما خلال فترة معينة:

النوع أ	٨٩	٩٠	٩٢	٨١	٧٦	٨٨	٨٤
النوع ب	٧٨	٩٣	٨١	٨٧	٨٩	٧١	٩٠
	٩٦	٨٢					

اختبر الفرض بأن النوعين لا يختلفان فى سعر البيع للوحدة بالقطاعى عند α = ٥٪ باستخدام اختبارين لابارامترى وآخر بارامترى.

١٨- فـيما يلـى المـدة (بالدقيقة) التى استغرقت فى إنجاز عملية ما فى أربعة فـروع مختلفة لأحد المصارف لعينة من العملاء عددهم ٦ عملاء فى كل حالة:

الفرع أ	٢٠	٢٥	٣٨	٣١	٢٣	٢٢
الفرع ب	١٨	١٩	١٢	٢١	٢٤	٩
الفرع جـ	١١	٨	١٠	٢٧	٢٥	١٤
الفرع د	١٤	١٧	٢٥	٢٨	٢١	١٥

حلـل نتائج هذه الدراسة مستخدماً تحليل التباين فى اتجاه واحد باستخدام الأسلوبين البارامترى واللابارامترى عند α = ٥٪ وناقش نتائجك.

١٩- فـيما يلـى عدد الكيلومترات / جالون التى استهلكت فى اختبار لمتوسط استهلاك الوقود لثلاثة أنواع من الوقود (أ ، ب ، جـ).

الفرع أ	٢٨	٢٣	٢٦	٣١	١٤	٢٩
الفرع ب	٢١	٣١	٣٢	١٩	٢٧	١٦
الفرع جـ	٢٤	١٧	٢١	٣١	٢٢	١٨

استخدم تحليل التباين فى اتجاه واحد (اختبار هـ) لتحليل نتائج هذه التجربة عند α = ٥٪ .

ثـم أعـد التحلـيل باسـتخدام تحلـيل التباين فى اتجاه واحد (الأسلوب البارامترى) عند α = ٥٪.

ثم ناقش النتائج التى تصل إليها فى الحالتين.

٢٠- لاختبار العلاقة بين الإنفاق على الأمن الصناعى (المنفق بالجنيه / عامل) ومعدل حوادث إصابات العمل فى ١١ مؤسسة صناعية فقد جمعت البيانات التالية:

المؤسسة	١	٢	٣	٤	٥	٦	٧
الإنفاق بالجنيه	٦٠	٤٥	٣٠	٢٠	٢٨	٤٢	٣٩
الحوادث	٣	٧	٦	٩	٧	٤	٨

تابع المؤسسة	٨	٩	١٠	١١
الإنفاق بالجنيه	٥٤	٤٨	٥٨	٢١
الحوادث	٢	٤	٣	٨

أوجـد قـيمة معـامل ارتـباط الرتب لسبيرمان بين الإنفاق على الأمن الصناعى وإصابات العمل. ثم اختبر معنوية هذا المعامل.

٢١- تعـرض البيانات التالية عدد القروض الجديدة التى قدمتها فرعين مختلفين من أفرع أحد المصارف وذلك خلال فترة الـ ١٦ يوماً السابقة:

فرع أ	فرع ب
٨	٨
٩	٦
٢	١٥
٦	٤
٧	١٥
٧	١٣
١٠	١٥
١٠	١٢

تابع

فرع أ	فرع ب
٦	١٣
٧	٦
١٨	٩
١١	١٤
١٠	٦
٣	١٥
١٥	٢
١٤	٣

اختبر الفرض العدمى $\mu_{أ} = \mu_{ب}$

$\mu_{أ} > \mu_{ب}$

عند مستوى المعنوية ٥٪ .

٢٢- إعتادت أحد الشركات أن تقدم هدايا مختلفة إلى بعض عملائها بداية كل عام ميلادى وذلك تحقيقاً لإستمرارية التعامل مع الشركة. وفى عام ما

قررت الشركة أن تقدم هدية معينة أ (نتيجة حائط تحمل اسم الشركة) إلى $\frac{١}{٢}$ عملائها وهدية من نوع آخر ب (طاقم أقلام يحمل الشركة) إلى الثلث الآخر وأما الثلث الأخير فلم تقدم له هدايا ثم قامت بتصنيف عملائها حسب نوع الهدايا ودرجة الإستجابة للتعامل مع الشركة وكان النحو التالى:

معدل الإستجابة	الهديـــــة		
	أ	ب	لا هدية
زاد	٤٥	٥٠	٥٥
ثابت	١٥	١٢	٢٠
نقص	٤	٣٨	٢٥

استخدم هذه البيانات فى اختبار أن نوع الهدية أو عدم تقديم هدية لم يؤثر علـــى درجـــة اســـتجابة العملاء مع الشركة عند α = ٥٪ وما هو نوع الاختـــبار ؟ ثـــم فسر ما تصل إليه نتائج وما هو القرار الإدارى الذى قد تنصح به فى حدود نتائج التحليل ؟

٢٣- تقدم أحد الفرق المسرحية عروضها ٣ مرات فى اليوم الواحد على مسرح صـــغير ، وفيما يلى عدد المتفرجين فى مرات العرض الثلاث فى يوم ما اختير عشوائياً.

العرض الأول	١٧٩	١٦٤	١٨١	١٧٦	١٧٣	
	١٧٨	١٦٤				
العرض الثانى	١٦٩	١٧٧	١٦٦	١٧١	١٨١	١٨٩
العرض الثالث	١٧٤	١٦٧	١٦٨	١٨٣	١٦٩	

اختبر الفرض (باختبار هـ) عند α = ٥٪ بعدم وجود اختلاف فى المتوسط الحقيقى لعدد المتفرجين فى مرات العرض الثلاث.

٢٤- فـيما يلى إجمالى المنفق من ميزانية الأسرة (س) فى سنة ما بالمائة جنيه ومـا أنفـق علـى المأكل (ص) فى ذات السنة بالمائة جنيه لأفراد عينة عشوائية مكونة من ١٠ أسر:

س	٢٠	٢٥	٣٠	٣٥	٤٠	٤٥	٥٠	٥٥	٦٠	٦٥
ص	١٤	١٥	١٧	١٦	٢٨	٢٠	٢١	٢٥	٢٤	٣٠

وإذا كانـت قيمة معامل الارتباط الخطى البسيط = ٠,٩٨٠ أحسب معامل ارتـباط الرتب لسبيرمان من هذه البيانات. ثم اختبر الرفض بأن معاملى الارتباط مقدراً بالطريقتين = صفر عند مستوى المعنوية ١٪ .

المراجع

أولاً : المراجع العربية :

١- الدكــتور / عــبد اللطــيف عــبد الفتاح أبو العلا : مدخل إلى الطرق الإحصائية – الطبعة الخامسة (١٩٩٣).

٢- الدكتور / عبد اللطيف عبد الفتاح أبو العلا : العينات وتصميم التجارب – الطبعة الثانية (١٩٩٦/١٩٩٧) جامعة المنصورة .

٣- الدكــتور / عــبد اللطــيف عبد الفتاح أبو العلا : التحليل الإحصائى ، جامعة المنصورة ، (١٩٩٧-١٩٩٨) .

٤- الدكــتور / عبد اللطيف عبد الفتاح أبو العلا ، الدكتور / سلطان محمد عــبد الحمــيد : أساســيات الإحصاء (١٩٩٩) مكتبة الجلاء الجديدة – المنصورة .

٥- الدكــتور / عــبد اللطيف عبد الفتاح أبو العلا ، الدكتور / محمد توفيق البلقيــنى ، الدكــتور / ســلطان محمد عبد الحميد : التحليل الإحصائى واســتخداماته فى العلوم التجارية والاجتماعية ، الجزء الثانى ، الطبعة الأولى (٢٠٠١/٢٠٠٢) .

٦- الإحصـــاء والاقتصاد القياسى- سلسلة شوم ترجمة دكتورة سعدية حافظ منتصر (١٩٨٣).

٧- الإحصاء فى الإدارة – كتاب مترجم دكتور عبد المرضى حامد عزام – دار المريخ ، المملكة العربية السعودية (١٩٩٦) .

ثانياً : المراجع الأجنبية :

1. Anderson, R.L. & Bancroft, T.A. : Statistical Theory In Research. McGraw-Hill Book Company, 1960.

2. C. R. Hicks Fundamental Concepts in the Design of Experiments, 3 rd. Fort Worth: Saunders College publishing, 1982

3. Canavos, J., C. and Miller, D.,M. (1999) : Modern Business Statistics, 7^{th} . ed.. International thomson Publishing Company, Duxbury Press, U.S.A

4. Cochrain, W.G., and Cox, G.M.: Experimental designs, John Wiley & Sons, 2^{nd}. ed., 1962.

5. Conover, W.J. : Practical Nonparametric Statistics, John Wiley & Sons Inc., 1971.

6. Conver, W.J., and Iman, R. L., Introduction To Modern Business Statistics, John Wiley &Jons Inc., 1983.

7. Dixon, Wilfrid & Massey, Frank Jr. : Introduction To Statistical Analysis, 4^{th}. ed., McGraw-Hil Book Company, 1983.

8. Fraser, D. A.: Statistics: An Introduction, John Wiley & Sons Inc., 1985

9. Freund, John, E & Williams, Frank, J. & Perles, Benjamin M: Elementary Business Statistics : The Modem Approach, 6^{th}. ed., Prentice-Hall, 1993.

10. Hoel, Poul G. & Jessen, Raymond J.: Basic Statistics For Business And Economics, 2^{nd}. ed., John Wiley & Sons Inc., 1977.

11. J. Neter, W. Wasserman, and M. Kutner, Applied Linear Statistical Models, 2nd ed. Homewood, IL: Richard D. Irwin, 1985.
12. Keller, Gerald & Warrak, Brain : Statistics for Management & Economics, 4th. ed., International Thomson Publishing Company, 1997.
13. Kohler, Heinz : Statistics For Business & Economics, 3rd . ed.. Harper Collins College Publishers, 1994.
14. L. Ott. An Introduction to Statistical Methods and Data Analysis, 4th ed. Belmont. CA: Duxbury Press, 1993.
15. Larsen R. J., and Marx, M. L.,: Statistics, Prentice-Hall International Editions, 1990.
16. Lindgren, B. W.: Statistical Theory, The Macmillan Company, 1962.
17. Maclave, James T. & Benson, George P. & Sincich, : Statistics For Business & Economics, 7th. ed., Prentice Hall International Editions, 1998.
18. Mann, Prem S. : Statistics For Business & Economics, John Wiley & Sons Inc., 1995.
19. Mason, R. L. and Lind, D. A. : Statistics Techniques in Business and Economics, 7th ed., Richard D. Irwin Inc., 1990.

20. Mood, Alexander M. & Grabill, Franklin A. : Introduction To The Theory Of Statistics, McGraw-Hill Book Company, 2nd . ed., 1963.

21. Ostle, O. : Statistics In Research, Iowa State University Press, 2nd. ed., 4th Printing 1969.

22. R. B. Miller and D. W. Wichern. Intermediate Business Statistics: Analysis of Variance, Regression and Time Series. New York: Holt, Rinehart & Winston. 1977.

23. R. D. Moen, T. W. Nolan and L. P. Provost. Improving Quality Through planned Experimentation. New York: Me Graw -Hill, Inc , 1991.

24. Sandy, Robert.: Statistics for Business and Economics, McGraw-Hill Book Company, International Editions, Statistics Series, 1990.

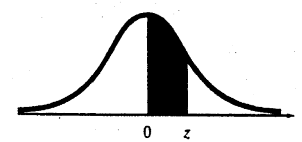
25. Selby, S. M., editor : Standard Mathematical Tables, XVI th. ed., The Chemical Rubber Company, 1958. 2S Sigel, Sidney : Nonparametric Statistics for behavioral Sciences, McGraw-Hill Book Company, 1956.

26. Snedecor, george & Cochran, William G. : Statistical Methods, 6th ed., The Lowa State University press, 1967.

27. Wonnacott, Thomas W., & Wonnacott, Ronald J. : Introductory Statistics For Business And Economics, 2nd. ed., John Wiley & Sons Inc., 1977.

محتويات الكتاب

فهرس الجداول الإحصائية



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Source: Abridged from Table I of A. Hald, *Statistical Tables and Formulas* (New York: Wiley), 1952. Reproduced by permission of A. Hald.

λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
.00	1.000000	**2.05**	.128735	**4.05**	.017422	**6.05**	.002358	**8.05**	.000319
.05	.951229	**2.10**	.122456	**4.10**	.016573	**6.10**	.002243	**8.10**	.000304
.10	.904837	**2.15**	.116484	**4.15**	.015764	**6.15**	.002133	**8.15**	.000289
.15	.860708	**2.20**	.110803	**4.20**	.014996	**6.20**	.002029	**8.20**	.000275
.20	.818731	**2.25**	.105399	**4.25**	.014264	**6.25**	.001930	**8.25**	.000261
.25	.778801	**2.30**	.100259	**4.30**	.013569	**6.30**	.001836	**8.30**	.000249
.30	.740818	**2.35**	.095369	**4.35**	.012907	**6.35**	.001747	**8.35**	.000236
.35	.704688	**2.40**	.090718	**4.40**	.012277	**6.40**	.001661	**8.40**	.000225
.40	.670320	**2.45**	.086294	**4.45**	.011679	**6.45**	.001581	**8.45**	.000214
.45	.637628	**2.50**	.082085	**4.50**	.011109	**6.50**	.001503	**8.50**	.000204
.50	.606531	**2.55**	.078082	**4.55**	.010567	**6.55**	.001430	**8.55**	.000194
.55	.576950	**2.60**	.074274	**4.60**	.010052	**6.60**	.001360	**8.60**	.000184
.60	.548812	**2.65**	.070651	**4.65**	.009562	**6.65**	.001294	**8.65**	.000175
.65	.522046	**2.70**	.067206	**4.70**	.009095	**6.70**	.001231	**8.70**	.000167
.70	.496585	**2.75**	.063928	**4.75**	.008652	**6.75**	.001171	**8.75**	.000158
.75	.472367	**2.80**	.060810	**4.80**	.008230	**6.80**	.001114	**8.80**	.000151
.80	.449329	**2.85**	.057844	**4.85**	.007828	**6.85**	.001059	**8.85**	.000143
.85	.427415	**2.90**	.055023	**4.90**	.007447	**6.90**	.001008	**8.90**	.000136
.90	.406570	**2.95**	.052340	**4.95**	.007083	**6.95**	.000959	**8.95**	.000130
.95	.386741	**3.00**	.049787	**5.00**	.006738	**7.00**	.000912	**9.00**	.000123
1.00	.367879	**3.05**	.047359	**5.05**	.006409	**7.05**	.000867	**9.05**	.000117
1.05	.349938	**3.10**	.045049	**5.10**	.006097	**7.10**	.000825	**9.10**	.000112
1.10	.332871	**3.15**	.042852	**5.15**	.005799	**7.15**	.000785	**9.15**	.000106
1.15	.316637	**3.20**	.040762	**5.20**	.005517	**7.20**	.000747	**9.20**	.000101
1.20	.301194	**3.25**	.038774	**5.25**	.005248	**7.25**	.000710	**9.25**	.000096
1.25	.286505	**3.30**	.036883	**5.30**	.004992	**7.30**	.000676	**9.30**	.000091
1.30	.272532	**3.35**	.035084	**5.35**	.004748	**7.35**	.000643	**9.35**	.000087
1.35	.259240	**3.40**	.033373	**5.40**	.004517	**7.40**	.000611	**9.40**	.000083
1.40	.246597	**3.45**	.031746	**5.45**	.004296	**7.45**	.000581	**9.45**	.000079
1.45	.234570	**3.50**	.030197	**5.50**	.004087	**7.50**	.000553	**9.50**	.000075
1.50	.223130	**3.55**	.028725	**5.55**	.003887	**7.55**	.000526	**9.55**	.000071
1.55	.212248	**3.60**	.027324	**5.60**	.003698	**7.60**	.000501	**9.60**	.000068
1.60	.201897	**3.65**	.025991	**5.65**	.003518	**7.65**	.000476	**9.65**	.000064
1.65	.192050	**3.70**	.024724	**5.70**	.003346	**7.70**	.000453	**9.70**	.000061
1.70	.182684	**3.75**	.023518	**5.75**	.003183	**7.75**	.000431	**9.75**	.000058
1.75	.173774	**3.80**	.022371	**5.80**	.003028	**7.80**	.000410	**9.80**	.000056
1.80	.165299	**3.85**	.021280	**5.85**	.002880	**7.85**	.000390	**9.85**	.000053
1.85	.157237	**3.90**	.020242	**5.90**	.002739	**7.90**	.000371	**9.90**	.000050
1.90	.149569	**3.95**	.019255	**5.95**	.002606	**7.95**	.000353	**9.95**	.000048
1.95	.142274	**4.00**	.018316	**6.00**	.002479	**8.00**	.000336	**10.00**	.000045
2.00	.135335								

جدول رقم (٣) .... توزيع ت

ν	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: This table is reproduced with the kind permission of the Trustees of Biometrika from E. S. Pearson and H. O. Hartley (eds.), *The Biometrika Tables for Statisticians*, Vol. 1, 3d ed., Biometrika, 1966.

جدول رقم (٤) .... توزيع ف ، $\alpha = 10$

NUMERATOR DEGREES OF FREEDOM

DENOMINATOR DEGREES OF FREEDOM ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

Source: From M. Merrington and C. M. Thompson. "Tables of Percentage Points of the Inverted Beta (F)-Distribution." *Biometrika*, 1943, 33, 73–88. Reproduced by permission of the *Biometrika* Trustees.

جدول رقم (٥) .... توزيع ف ، $\alpha = .05$

NUMERATOR DEGREES OF FREEDOM (ν_1) / DENOMINATOR DEGREES OF FREEDOM (ν_2)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.77	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Source: From M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta (*F*)-Distribution." *Biometrika*, 1943, 33, 73–88.

جدول رقم (٦) .... توزيع ف ، $\alpha = .025$

NUMERATOR DEGREES OF FREEDOM (ν_1) / **DENOMINATOR DEGREES OF FREEDOM** (ν_2)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1,001	1,006	1,010	1,014	1,018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

Source: From M. Merrington and C. M. Thompson. "Tables of Percentage Points of the Inverted Beta (F)-Distribution," *Biometrika*, 1943, 33, 73–88.

جدول رقم (٧) .... توزيع ف ، $\alpha = .01$

NUMERATOR DEGREES OF FREEDOM

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4,052	4,999.5	5,403	5,625	5,764	5,859	5,928	5,982	6,022	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

DENOMINATOR DEGREES OF FREEDOM

Source: From M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta (*F*)-Distribution," *Biometrika*, 1943, 33, 73–88.

جدول رقم (٨) ....اختبار ولكوكسن : العينات المستقلة

Test statistic is the rank sum associated with the smaller sample (if equal sample sizes, either rank sum can be used).

a. $\alpha = .025$ one-tailed; $\alpha = .05$ two-tailed

n_2 \ n_1	3		4		5		6		7		8		9		10	
	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U
3	5	16	6	18	6	21	7	23	7	26	8	28	8	31	9	33
4	6	18	11	25	12	28	12	32	13	35	14	38	15	41	16	44
5	6	21	12	28	18	37	19	41	20	45	21	49	22	53	24	56
6	7	23	12	32	19	41	26	52	28	56	29	61	31	65	32	70
7	7	26	13	35	20	45	28	56	37	68	39	73	41	78	43	83
8	8	28	14	38	21	49	29	61	39	73	49	87	51	93	54	98
9	8	31	15	41	22	53	31	65	41	78	51	93	63	108	66	114
10	9	33	16	44	24	56	32	70	43	83	54	98	66	114	79	131

b. $\alpha = .05$ one-tailed; $\alpha = .10$ two-tailed

n_2 \ n_1	3		4		5		6		7		8		9		10	
	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U
3	6	15	7	17	7	20	8	22	9	24	9	27	10	29	11	31
4	7	17	12	24	13	27	14	30	15	33	16	36	17	39	18	42
5	7	20	13	27	19	36	20	40	22	43	24	46	25	50	26	54
6	8	22	14	30	20	40	28	50	30	54	32	58	33	63	35	67
7	9	24	15	33	22	43	30	54	39	66	41	71	43	76	46	80
8	9	27	16	36	24	46	32	58	41	71	52	84	54	90	57	95
9	10	29	17	39	25	50	33	63	43	76	54	90	66	105	69	111
10	11	31	18	42	26	54	35	67	46	80	57	95	69	111	83	127

Source: From F. Wilcoxon and R. A. Wilcox, "Some Rapid Approximate Statistical Procedures," 1964, 20–23. Courtesy of Lederle Laboratories Division of American Cyanamid Company, Madison, NJ.

جدول رقم (٩) .... اختبار ولكوكسن : العينات المستقلة

One-Tailed	Two-Tailed	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$\alpha = .05$	$\alpha = .10$	1	2	4	6	8	11
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8
$\alpha = .01$	$\alpha = .02$			0	2	3	5
$\alpha = .005$	$\alpha = .01$				0	2	3
		$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$\alpha = .05$	$\alpha = .10$	14	17	21	26	30	36
$\alpha = .025$	$\alpha = .05$	11	14	17	21	25	30
$\alpha = .01$	$\alpha = .02$	7	10	13	16	20	24
$\alpha = .005$	$\alpha = .01$	5	7	10	13	16	19
		$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$\alpha = .05$	$\alpha = .10$	41	47	54	60	68	75
$\alpha = .025$	$\alpha = .05$	35	40	46	52	59	66
$\alpha = .01$	$\alpha = .02$	28	33	38	43	49	56
$\alpha = .005$	$\alpha = .01$	23	28	32	37	43	49
		$n = 23$	$n = 24$	$n = 25$	$n = 26$	$n = 27$	$n = 28$
$\alpha = .05$	$\alpha = .10$	83	92	101	110	120	130
$\alpha = .025$	$\alpha = .05$	73	81	90	98	107	117
$\alpha = .01$	$\alpha = .02$	62	69	77	85	93	102
$\alpha = .005$	$\alpha = .01$	55	61	68	76	84	92
		$n = 29$	$n = 30$	$n = 31$	$n = 32$	$n = 33$	$n = 34$
$\alpha = .05$	$\alpha = .10$	141	152	163	175	188	201
$\alpha = .025$	$\alpha = .05$	127	137	148	159	171	183
$\alpha = .01$	$\alpha = .02$	111	120	130	141	151	162
$\alpha = .005$	$\alpha = .01$	100	109	118	128	138	149
		$n = 35$	$n = 36$	$n = 37$	$n = 38$	$n = 39$	
$\alpha = .05$	$\alpha = .10$	214	228	242	256	271	
$\alpha = .025$	$\alpha = .05$	195	208	222	235	250	
$\alpha = .01$	$\alpha = .02$	174	186	198	211	224	
$\alpha = .005$	$\alpha = .01$	160	171	183	195	208	
		$n = 40$	$n = 41$	$n = 42$	$n = 43$	$n = 44$	$n = 45$
$\alpha = .05$	$\alpha = .10$	287	303	319	336	353	371
$\alpha = .025$	$\alpha = .05$	264	279	295	311	327	344
$\alpha = .01$	$\alpha = .02$	238	252	267	281	297	313
$\alpha = .005$	$\alpha = .01$	221	234	248	262	277	292
		$n = 46$	$n = 47$	$n = 48$	$n = 49$	$n = 50$	
$\alpha = .05$	$\alpha = .10$	389	408	427	446	466	
$\alpha = .025$	$\alpha = .05$	361	379	397	415	434	
$\alpha = .01$	$\alpha = .02$	329	345	362	380	398	
$\alpha = .005$	$\alpha = .01$	307	323	339	356	373	

Source: From F. Wilcoxon and R. A. Wilcox, "Some Rapid Approximate Statistical Procedures," 1964, p. 28. Courtesy of Lederle Laboratories Division of American Cyanamid Company, Madison, NJ.

جدول رقم (١٠) كا٢

Degrees of Freedom	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	.0000393	.0001571	.0009821	.0039321	.0157908	2.70554	3.84146	5.02389	6.63490	7.87944
2	.0100251	.0201007	.0506356	.102587	.210720	4.60517	5.99147	7.37776	9.21034	10.5966
3	.0717212	.114832	.215795	.351846	.584375	6.25139	7.81473	9.34840	11.3449	12.8381
4	.206990	.297110	.484419	.710721	1.063623	7.77944	9.48773	11.1433	13.2767	14.8602
5	.411740	.554300	.831211	1.145476	1.61031	9.23635	11.0705	12.8325	15.0863	16.7496
6	.675727	.872085	1.237347	1.63539	2.20413	10.6446	12.5916	14.4494	16.8119	18.5476
7	.989265	1.239043	1.68987	2.16735	2.83311	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.344419	1.646482	2.17973	2.73264	3.48954	13.3616	15.5073	17.5346	20.0902	21.9550
9	1.734926	2.087912	2.70039	3.32511	4.16816	14.6837	16.9190	19.0228	21.6660	23.5893
10	2.15585	2.55821	3.24697	3.94030	4.86518	15.9871	18.3070	20.4831	23.2093	25.1882
11	2.60321	3.05347	3.81575	4.57481	5.57779	17.2750	19.6751	21.9200	24.7250	26.7569
12	3.07382	3.57056	4.40379	5.22603	6.30380	18.5494	21.0261	23.3367	26.2170	28.2995
13	3.56503	4.10691	5.00874	5.89186	7.04150	19.8119	22.3621	24.7356	27.6883	29.8194
14	4.07468	4.66043	5.62872	6.57063	7.78953	21.0642	23.6848	26.1190	29.1413	31.3193
15	4.60094	5.22935	6.26214	7.26094	8.54675	22.3072	24.9958	27.4884	30.5779	32.8013
16	5.14224	5.81221	6.90766	7.96164	9.31223	23.5418	26.2962	28.8454	31.9999	34.2672
17	5.69724	6.40776	7.56418	8.67176	10.0852	24.7690	27.5871	30.1910	33.4087	35.7185
18	6.26481	7.01491	8.23075	9.39046	10.8649	25.9894	28.8693	31.5264	34.8053	37.1564
19	6.84398	7.63273	8.90655	10.1170	11.6509	27.2036	30.1435	32.8523	36.1908	38.5822
20	7.43386	8.26040	9.59083	10.8508	12.4426	28.4120	31.4104	34.1696	37.5662	39.9968
21	8.03366	8.89720	10.28293	11.5913	13.2396	29.6151	32.6705	35.4789	38.9321	41.4010
22	8.64272	9.54249	10.9823	12.3380	14.0415	30.8133	33.9244	36.7807	40.2894	42.7956
23	9.26042	10.19567	11.6885	13.0905	14.8479	32.0069	35.1725	38.0757	41.6384	44.1813
24	9.88623	10.8564	12.4011	13.8484	15.6587	33.1963	36.4151	39.3641	42.9798	45.5585
25	10.5197	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3141	46.9278
26	11.1603	12.1981	13.8439	15.3791	17.2919	35.5631	38.8852	41.9232	45.6417	48.2899
27	11.8076	12.8786	14.5733	16.1513	18.1138	36.7412	40.1133	43.1944	46.9630	49.6449
28	12.4613	13.5648	15.3079	16.9279	18.9392	37.9159	41.3372	44.4607	48.2782	50.9933
29	13.1211	14.2565	16.0471	17.7083	19.7677	39.0875	42.5569	45.7222	49.5879	52.3356
30	13.7867	14.9535	16.7908	18.4926	20.5992	40.2560	43.7729	46.9792	50.8922	53.6720
40	20.7065	22.1643	24.4331	26.5093	29.0505	51.8050	55.7585	59.3417	63.6907	66.7659
50	27.9907	29.7067	32.3574	34.7642	37.6886	63.1671	67.5048	71.4202	76.1539	79.4900
60	35.5346	37.4848	40.4817	43.1879	46.4589	74.3970	79.0819	83.2976	88.3794	91.9517
70	43.2752	45.4418	48.7576	51.7393	55.3290	85.5271	90.5312	95.0231	100.425	104.215
80	51.1720	53.5400	57.1532	60.3915	64.2778	96.5782	101.879	106.629	112.329	116.321
90	59.1963	61.7541	65.6466	69.1260	73.2912	107.565	113.145	118.136	124.116	128.299
100	67.3276	70.0648	74.2219	77.9295	82.3581	118.498	124.342	129.561	135.807	140.169

Source: From C. M. Thompson, "Tables of the Percentage Points of the χ^2-Distribution," *Biometrika*, 1941, 32, 188–189. Reproduced by permission of the *Biometrika* Trustees.

جدول رقم (۱۱) ....احصاء دیربن ـــ واطسون $\alpha = .05$

n	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	.95	1.54	.82	1.75	.69	1.97	.56	2.21
16	1.10	1.37	.98	1.54	.86	1.73	.74	1.93	.62	2.15
17	1.13	1.38	1.02	1.54	.90	1.71	.78	1.90	.67	2.10
18	1.16	1.39	1.05	1.53	.93	1.69	.92	1.87	.71	2.06
19	1.18	1.40	1.08	1.53	.97	1.68	.86	1.85	.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	.90	1.83	.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	.93	1.81	.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	.96	1.80	.96	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	.99	1.79	.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

Source: From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika*, 1951, 30, 159–178. Reproduced by permission of the *Biometrika* Trustees.

جدول رقم (١٢) .... احصاء ديربن — واطسون $\alpha = .01$

	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	.81	1.07	.70	1.25	.59	1.46	.49	1.70	.39	1.96
16	.84	1.09	.74	1.25	.63	1.44	.53	1.66	.44	1.90
17	.87	1.10	.77	1.25	.67	1.43	.57	1.3	.48	1.85
18	.90	1.12	.80	1.26	.71	1.42	.61	1.60	.52	1.80
19	.93	1.13	.83	1.26	.74	1.41	.65	1.58	.56	1.77
20	.95	1.15	.86	1.27	.77	1.41	.68	1.57	.60	1.74
21	.97	1.16	.89	1.27	.80	1.41	.72	1.55	.63	1.71
22	1.00	1.17	.91	1.28	.83	1.40	.75	1.54	.66	1.69
23	1.02	1.19	.94	1.29	.86	1.40	.77	1.53	.70	1.67
24	1.04	1.20	.96	1.30	.88	1.41	.80	1.53	.72	1.66
25	1.05	1.21	.98	1.30	.90	1.41	.83	1.52	.75	1.65
26	1.07	1.22	1.00	1.31	.93	1.41	.85	1.52	.78	1.64
27	1.09	1.23	1.02	1.32	.95	1.41	.88	1.51	.81	1.63
28	1.10	1.24	1.04	1.32	.97	1.41	.90	1.51	.83	1.62
29	1.12	1.25	1.05	1.33	.99	1.42	.92	1.51	.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	.94	1.51	.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	.96	1.51	.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	.98	1.51	.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	.95	1.59
35	1.19	1.31	1.14	1.27	1.08	1.44	1.03	1.51	.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

Source: From J. Durbin and G. S. Watson. "Testing for Serial Correlation in Least Squares Regression. II." *Biometrika.* 1951, 30. 159–178. Reproduced by permission of the *Biometrika* Trustees.

Non-Parametric Statistics

PROBABILITIES ASSOCIATED WITH VALUES AS LARGE AS OBSERVED VALUES OF H IN THE KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE BY RANKS

Sample sizes			H	p
n_1	n_2	n_3		
5	2	2	6.5333	.008
			6.1333	.013
			5.1600	.034
			5.0400	.056
			4.3733	.090
			4.2933	.122
5	3	1	6.4000	.012
			4.9600	.048
			4.8711	.052
			4.0178	.095
			3.8400	.123
5	3	2	6.9091	.009
			6.8218	.010
			5.2509	.049
			5.1055	.052
			4.6509	.091
			4.4945	.101
5	3	3	7.0788	.009
			6.9818	.011
			5.6485	.049
			5.5152	.051
			4.5333	.097
			4.4121	.109
5	4	1	6.9545	.008
			6.8400	.011
			4.9855	.044
			4.8600	.056
			3.9873	.098
			3.9600	.102
5	4	2	7.2045	.009
			7.1182	.010
			5.2727	.049
			5.2682	.050
			4.5409	.098
			4.5182	.101
5	4	3	7.4449	.010
			7.3949	.011
			5.6564	.049
			5.6308	.050
			4.5487	.099
			4.5231	.103
	4	4	7.7604	.009
			7.7440	.011
			5.6571	.049
			5.6176	.050
			4.6187	.100
			4.5527	.102
5	5	1	7.3091	.009
			6.8364	.011
			5.1273	.046
			4.9091	.053
			4.1091	.086
			4.0364	.105
5	5	2	7.3385	.010
			7.2692	.010
			5.3385	.047
			5.2462	.051
			4.6231	.097
			4.5077	.100
5	5	3	7.5780	.010
			7.5429	.010
			5.7055	.046
			5.6264	.051
			4.5451	.100
			4.5363	.102
5	5	4	7.8229	.010
			7.7914	.010
			5.6657	.049
			5.6429	.050
			4.5229	.099
			4.5200	.101
5	5	5	8.0000	.009
			7.9800	.010
			5.7800	.049
			5.6600	.051
			4.5600	.100
			4.5000	.102

Non-Parametric Statistics

PROBABILITIES ASSOCIATED WITH VALUES AS LARGE AS OBSERVED VALUES OF H IN THE KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE BY RANKS

Sample sizes			H	p
n_1	n_2	n_3		
2	1	1	2.7000	.500
2	2	1	3.6000	.200
2	2	2	4.5714	.067
			3.7143	.200
3	1	1	3.2000	.300
3	2	1	4.2857	.100
			3.8571	.133
3	2	2	5.3572	.029
			4.7143	.048
			4.5000	.067
			4.4643	.105
3	3	1	5.1429	.043
			4.5714	.100
			4.0000	.129
3	3	2	6.2500	.011
			5.3611	.032
			5.1389	.061
			4.5556	.100
			4.2500	.121
3	3	3	7.2000	.004
			6.4889	.011
			5.6889	.029
			5.6000	.050
			5.0667	.086
			4.6222	.100
4	1	1	3.5714	.200
4	2	1	4.8214	.057
			4.5000	.076
			4.0179	.114
4	2	2	6.0000	.014
			5.3333	.033
			5.1250	.052
			4.4583	.100
			4.1667	.105
4	3	1	5.8333	.021
			5.2083	.050
			5.0000	.057
			4.0556	.093
			3.8889	.129
4	3	2	6.4444	.008
			6.3000	.011
			5.4444	.046
			5.4000	.051
			4.5111	.098
			4.4444	.102
4	3	3	6.7455	.010
			6.7091	.013
			5.7909	.046
			5.7273	.050
			4.7091	.092
			4.7000	.101
4	4	1	6.6667	.010
			6.1667	.022
			4.9667	.048
			4.8667	.054
			4.1667	.082
			4.0667	.102
4	4	2	7.0364	.006
			6.8727	.011
			5.4545	.046
			5.2364	.052
			4.5545	.098
			4.4455	.103
4	4	3	7.1439	.010
			7.1364	.011
			5.5985	.049
			5.5758	.051
			4.5455	.099
			4.4773	.102
4	4	4	7.6538	.008
			7.5385	.011
			5.6923	.049
			5.6538	.054
			4.6539	.097
			4.5001	.104
5	1	1	3.8571	.143
5	2	1	5.2500	.036
			5.0000	.048
			4.4500	.071
			4.2000	.095
			4.0500	.119